Seismic risk assessment of complex urban critical infrastructure networks

by

Indranil Kongar

A thesis submitted for the degree of Doctor of Engineering
Department of Civil, Environmental and Geomatic Engineering
University College London

May 2017

Declaration

I, Indranil Kongar, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Date:

Abstract

Existing practice on seismic risk assessment of critical infrastructure systems is reviewed in terms of exposure, hazard, fragility, performance and interdependencies and a framework is proposed for seismic risk assessment model development for the insurance sector. The application of the framework is demonstrated with the electric power and water supply systems in Christchurch, New Zealand. This includes the development of the first ground motion residual spatial correlation model for the region and a simplified method for predicting the occurrence of liquefaction. Empirical data from the 2010-11 Canterbury earthquake sequence are used to derive new fragility functions for substations, buried cables, wells, pumping stations and pipes, in terms of both ground shaking and permanent ground deformation. The model is tested against performance observations from the February 22nd 2011 M_W 6.2 Christchurch earthquake and achieves reasonable results when interdependencies between substations and pumps are modelled by nodal analysis with mapped substation supply zones rather than proximity rules. The model is applied to construct a future risk projection and in a 10,000-year stochastic catalogue, the electric power network exhibits high reliability with performance loss in only 2% of events. The water supply system is less reliable, due to the effect of ground shaking and liquefaction on pipes and the effect of power loss on the functionality of pumps, which is shown to increase disconnections by up 30%. A new metric, the interdependency index, is proposed to measure the degree of the dependency of the water supply system on electric power. It is adapted from the Leontief input-output method for infrastructure interdependency modelling and makes use of the system performance results acquired from the future risk projection, by assuming a linear relationship between the change in performance of the water supply system due to power loss and the performance of the electric power network.

Acknowledgments

The research presented in this thesis was co-funded by the Engineering and Physical Sciences Research Council in the United Kingdom and the Willis Research Network. Firstly I would like to thank my supervisor, Professor Tiziana Rossetto, for providing me with the opportunity to research such an interesting and challenging topic and for providing valuable technical support and guidance over the last five years. The Christchurch (New Zealand) case study presented in this thesis would not have been possible without Dr. Sonia Giovinazzi at the University of Canterbury (New Zealand) and so I extend a special thanks to her for her willingness to collaborate and share data and for allowing me the opportunity to conduct research under her supervision in New Zealand. I am indebted to all those in Christchurch who provided me with the datasets needed to apply my case study, including Dave Brannigan, Peter Elliott and Ricki-Lee Teague at Orion New Zealand, Sjoerd van Ballegooy at Tonkin and Taylor, Nicola Litchfield at GNS Science and Dr. Matthew Hughes at the University of Canterbury. I am also grateful for the technical and professional guidance I have received from Dr. Ioanna Ioannou and Enrica Verrucci at University College London and from Dr. John Alarcon and Dr. Myrto Papaspiliou at Willis Towers Watson. I would like to give a big thank you to all of my friends, but especially Patrick, Marco, Pierre and Josh for making the office a more bearable place to be for all those hours. Finally, I would like to thank my family for all their love and moral support, but especially Cat, with whom I have shared this experience most closely, for her tolerance, understanding and unceasing emotional encouragement.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AC	Asbestos cement
AIC	Akaike Information Criterion
AIR	Applied Insurance Research
ALA	American Lifelines Alliance
AMP	Asset management plan
ATC	Applied Technology Council
AUC	Area under receiver operating characteristic curve
CBD	Central business district
CCC	Christchurch City Council
CERA	Canterbury Earthquake Recovery Authority
CI	Confidence interval
CLDI	Cement lined ductile iron
CLS	Cement lined steel
CPT	Cone penetration test
CSR	Cyclic stress ratio
CTI	Compound topographic index
DEM	Digital elevation model
df	Degrees of freedom
DHS	Department of Homeland Security
DI	Ductile iron
DSHA	Deterministic seismic hazard assessment
ECan	Environment Canterbury Regional Council
EP	Exceedance probability
EQC	Earthquake Commission
EPN	Electric power network
FEMA	Federal Emergency Management Agency
FN	False negatives

FP	False positives
FPR	False positive rate
GALV	Galvanised steel
GDP	Gross domestic product
GEM	Global Earthquake Model
GIS	Geographic information system
GMPE	Ground motion prediction equation
GXP	Grid exit point
HDPE	High density polyethylene
IM	Intensity measure
LPI	Liquefaction potential index
MCC	Matthew's correlation coefficient
MDPE	Medium density polyethylene
MPVC	Modified polyvinyl chloride
ND	Normalised distance
NSHM	National Seismic hazard Model
NZD	New Zealand dollars
NZGD	New Zealand Geotechnical Database
PEER	Pacific Earthquake Engineering Research Center
PGA	Peak ground acceleration
PGD	Peak ground displacement
PGDf	Permanent ground deformation
$PGDf_{GEOM}$	Geometric mean of horizontal and vertical permanent ground deformations due to liquefaction
$PGDf_{H}$	Horizontal permanent ground deformation due to liquefaction
$PGDf_{LAND}$	Permanent ground deformation due to landslide
$PGDf_{MAX}$	Maximum of horizontal and vertical permanent ground deformations due to liquefaction
$PGDf_{RUP}$	Permanent ground deformation due to fault rupture
$PGDf_{RUPMAX}$	Maximum permanent ground deformation due to fault rupture

$PGDf_V$	Vertical permanent ground deformation due to liquefaction
$PGDf_{VECT}$	Vector mean of horizontal and vertical permanent ground deformations due to liquefaction
PGV	Peak ground velocity
PILCA	Paper insulated lead covered armoured
PSHA	Probabilistic seismic hazard assessment
PVC	Polyvinyl chloride
RMS	Risk Management Solutions
ROC	Receiver operating characteristic
RR	Repair rate
SA	Spectral acceleration
SD	Spectral displacement
SE	Standard error
SPT	Standard penetration test
SV	Spectral velocity
TCLEE	Technical Council of Lifeline Earthquake Engineering
TN	True negatives
TNR	True negative rate
TP	True positives
TPR	True positive rate
UPVC	Unplasticised polyvinyl chloride
USD	United States dollars
USGS	United States Geological Survey
XLPE	Cross-linked polyethylene
WSS	Water supply system

List of Symbols

Symbol	Meaning
A	Matrix of system interdependencies in Leontief input-output model
a_c	Critical acceleration in calculation of landslide-induced permanent ground deformation (Saygili and Rathje, 2008)
a_{dim}	Source dimension of earthquake interaction zone
a_{GR}, b_{GR}	Coefficients of Gutenberg-Richter relationship
a_{IR}, b_{IR}, c_{IR}	Coefficients of interdependency response regression model
a_{max}	Peak horizontal ground acceleration
a_{RR}, b_{RR}	Coefficients of repair rate function based on power relationship
a_{SC}	Sill in relationship between semivariogram and correlation
b_f	Local bandwidth factor for Gaussian kernel function (seismic source model)
b_{JB}	Range of spatial correlation model (Jayaram and Baker, 2009)
C	Matrix of site to site correlations of simulated ground motion intra-event residuals
c	Perturbation vector in Leontief input-output model
c_1, c_3, c_m, $c_{\gamma 2}$	Coefficients of Chiou and Youngs (2008) and Chiou et al. (2010) PGV prediction models
ct	Number of years for which earthquake catalogue is complete for a specified magnitude
d_{cell}	Distance between cells for development seismic source model
d_{corr}	Correlation distance for Gaussian kernel function (seismic source model)
d_i	The *i*th fragility function damage state
d_w	Depth to ground water
F^*	Integrand for calculation of liquefaction potential index
F_{HW}	Hanging wall flag
F_{NM}	Normal faulting flag
F_{RV}	Reverse faulting flag
FS	Factor of safety against liquefaction
FS^*	Unbiased factor of safety against liquefaction
g	Acceleration of gravity (9.81 m/s^2)
h	Separation distance between two ground motion observation sites

Symbol	Meaning
I	Identity matrix
I_{21}	Interdependency index quantifying the effect of loss of performance in system 2 on the performance of system 1 (i.e. the dependency of system 1 on system 2)
IM_{di}	Median value of ground motion IM for the *i*th damage state
IR_{WSS}	Interdependency response in the water supply system
$I_{WSS\|EPN}$	Interdependency index quantifying the dependency of the water supply system on the electric power network
k	Number of independent variables in interdependency response model
K_M	Moment magnitude correction factor
K_W	Groundwater correction factor
K_Δ	Displacement correction factor
L	Lower triangular matrix determined by Cholesky factorisation of **C**
L	Likelihood estimate
L_{cable}	Length of buried cable
L_{min}	Minimum length of cable/pipe required for meaningful repair rate analysis
L_{pipe}	Length of pipe
L_{rup}	Length of fault rupture
M	Magnitude
M_0	First moment of earthquake
M_{cutoff}	Maximum magnitude for a source zone
M_L	Local magnitude
M_{max}	Maximum magnitude observed in an aftershock sequence
M_{min}	Magnitude for which earthquake catalogue is complete
M_S	Surface wave magnitude
MSF	Magnitude scaling factor for cyclic resistance ratio
M_W	Moment magnitude
MWF	Magnitude weighting factor
N_{GR}	Annual frequency of events greater than or equal to specified magnitude in Gutenberg-Richter relationship
n_{IR}	Number of observations used in regression to derive interdependency response model

Symbol	Meaning
NL	Non-linear response term for intra-event residual standard error model
NL_0	Value of NL calculated when η is zero
n_{LAND}	Number of cycles in calculation of landslide-induced permanent ground deformation (Hazus)
n_{RR}	Number of observations used in regression to derive repair rate function
N_{SC}	Number of pairs of sites used to estimate semivariogram
P_a	Reference stress for cyclic resistance ratio
P_{ml}	Proportion of map unit susceptible to liquefaction
R_{cable}	Number of buried cable repairs
r_d	Shear stress reduction coefficient
R_{earth}	Radius of the Earth (6,371 km)
R_f	Horizontal distance to surface projection of fault rupture
R_{JB}	The horizontal distance to the surface projection of the fault rupture plane
RR_{obs}	Observed repair rate
R_{rup}	The slant distance to the closest point on fault rupture plane
R'_{rup}	The in-plane rupture distance
R_X	Distance from the surface projection of the up-dip edge of the fault plane, measured perpendicular to the fault strike (positive in the down-dip direction)
R_Y	The closest distance to the rupture surface projection measured parallel to the strike
SP	System performance
SP_{EPN}	Performance of the electric power network
SP_{WSS}	Performance of the water supply system after interdependencies are accounted for
SP'_{WSS}	Performance of the water supply system before interdependencies are accounted for
SS_{res}	Residual sum of squares in regression
S_{yx}	Standard error or repair rate estimates from regression model
t	Time between two successive earthquakes
t_{crit}	Critical t-statistic with (n_{RR}-2) degrees of freedom and significance level of α_{RR}
T_d	Time between first and last occurrence of 0.05 g in accelerogram
u	Location of a ground motion observation sites

Symbol	Meaning
U_1	Uniformly distributed random variable (between 0 and 1) for seismic event generation
V^*_{S1}	Upper limit of V_{S1}
V_S	Shear wave velocity
$V_{S(0\text{-}10)}$	Average shear wave velocity in top 10 m of soil
$V_{S(10\text{-}20)}$	Average shear wave velocity for soil layers between 10 m and 20 m depth
V_{S1}	Stress corrected shear wave velocity
V_{S30}	Average shear wave velocity in top 30 m
W	Width of fault rupture
$w_{(i,EPN)}$	Weight (0 or 1) assigned to meshblock signifying local functionality of electric power network
$w_{(i,WSS)}$	Weight (0 or 1) assigned to meshblock signifying local functionality of water supply system
X	Linear function of predictor variables for global (X_G), regional (X_R) and local (X_L) liquefaction prediction models by Zhu et al. (2015)
x	Vector of system inoperabilities in Leontief input-output model
Y	Vector of spatially correlated ground motion intra-event residuals
Y^*	Probit link function for generalised linear model forecasting liquefaction probability from liquefaction potential index
Z	Vector of uncorrelated ground motion intra-event residuals
z	Depth of soil layer for calculation of cyclic stress ratio
$Z_{1.0}$	Depth to a shear wave velocity of 1.0 km/s
Z_{HYP}	Hypocentral depth
Z_{TOR}	Depth to top of rupture
α	Source-to-site azimuth
α_B	Coefficient of Boore et al. (2003) spatial correlation model
$\boldsymbol{\alpha}_{FRAG}$	Vector of observed intensity measures for Bayesian updating of fragility functions
α_{GA}, β_{GA}, κ_{GA}	Coefficients of Goda and Atkinson (2009) spatial correlation model
α_{RR}	Significance level repair rate regression model confidence interval
α_{WT}	Coefficient of Wang and Takada (2005) spatial correlation model
β_0, β_1	Coefficients of probit link function for generalised linear model forecasting liquefaction probability from liquefaction potential index

Symbol	Meaning
β_{di}	Dispersion of ground motion IM for the *i*th damage state
$\boldsymbol{\beta}_{FRAG}$	Vector of observed damage states for Bayesian updating of fragility functions
$\boldsymbol{\beta}_{GMPE}$	Vector of coefficients for ground motion prediction equation
β_{IR}	Slope of interdependency response regression model for Z-test
γ	Semivariogram of normalised intra-event residuals at a specified separation distance
δ	Dip
Δ_{SP}	Absolute change in water supply system performance in sensitivity test
ε	Intra-event residual for ground motion prediction equation
ε_{LAND}	Normally distributed random variable for standard deviation of Saygili and Rathje (2008) landslide-induced permanent ground deformation model
ε_T	Total residual for ground motion prediction equation
η	Inter-event residual for ground motion prediction equation
η	Random effects term for calculating non-linear response term for intra-event residual standard error model
$\boldsymbol{\theta}_{FRAG}$	Vector of proposed parameter values for Bayesian updating of fragility functions
$\boldsymbol{\theta}_{GMPE}$	Vector of predictor variables for ground motion prediction equation
λ	Poisson rate parameter for seismic event generation
$\boldsymbol{\mu}$	Vector of mean ground motion intra-event residuals
μ	Mean value of seismic source zone parameter across source grid
ρ	Correlation of normalised intra-event residuals at a specified separation distance
$\rho_{PGA,PGV}$	Correlation between total residuals of PGA and PGV
ρ_{WSS}	Proportion of additional water supply system customers disconnected due to physical damage in sensitivity test that are also disconnected due to power failure
σ	Standard error of intra-event residual for ground motion prediction equation
σ_{NLo}	Value of σ_T calculated when η is zero
σ_T	Standard error of total residual for ground motion prediction equation
σ_v	Total overburden stress
σ'_v	Effective overburden stress
τ	Standard error of inter-event residual for ground motion prediction equation
τ_{IZ}	Temporal component of earthquake interaction zone

1. Introduction

Earthquakes are amongst the most devastating of natural hazards, giving rise to six of the ten largest disasters since 1900, in terms of number of casualties (Guha-Sapir et al., 2015). They also result in significant economic impacts for affected regions. Between 1900 and 2011, it is estimated that total losses due to earthquakes exceeded $4 trillion USD (adjusted to present day values) (Daniell et al., 2012). There have been nine earthquakes in which total losses have amounted to more than half of the annual gross domestic product of the affected country (Daniell et al., 2011). The importance of critical infrastructure performance and business interruption in seismic events is exemplified by the 1995 earthquake in Kobe, Japan. It is estimated that nearly 30% of direct economic losses derived from infrastructure damage rather than from damage to buildings, and $100 billion USD worth of losses were generated by subsequent business interruption (Chang, 1996).

In middle and high income countries, insurance can play a vital role in mitigating the economic impacts of earthquakes. The United States and Japan are the two largest earthquake insurance markets (Coburn and Spence, 2002) but the benefit of insurance is perhaps best illustrated by the sequence of earthquakes that struck the Canterbury region of New Zealand between 2010 and 2011. These generated losses in excess of $20 billion USD between them and approximately 80% of these were insured, far higher in percentage terms than other earthquakes that occurred around the same time (King et al., 2014). However, currently the risks arising from damage to complex urban infrastructure systems are poorly understood by the insurance industry. The focus of this research, which is funded by the reinsurance advisors, Willis Re, is to help meet the demand within the industry to better understand the implications of these risks, with specific reference to earthquakes. The remainder of this chapter provides some background on catastrophe risk modelling in the insurance sector and a summary of key conceptual issues relating to the risk assessment of critical infrastructure systems. It then goes on to describe how this research has been approached and summarises the structure of the remainder of this thesis.

1.1 Catastrophe insurance

Insurance is a process of risk transfer. Insurers take on risk in return for a premium payment and agree to pay out to the policy holders if a loss-causing event occurs. Insurers then seek to transfer some of their risk to reinsurers, who in turn can seek to transfer risk to the private investment market, by selling catastrophe bonds – bonds that mature and pay back to the purchaser only if no defined natural catastrophe occurs over the bond period (Woo, 2002). Although there are examples of government-backed insurance schemes (e.g. the Earthquake Commission in New Zealand, Flood Re in the United Kingdom), insurers are generally commercial entities. Their business relies on pooling together premiums from policy holders whose risk of requiring compensation at the same time is statistically independent. In the case of natural hazards this is most easily achieved by ensuring geographical diversity. The aim of an insurer is to set premiums such that there is a low probability in a given year that the claims paid out exceed the sum of the premiums paid in and their cash reserves (Brillinger, 1993). This process, known as risk pricing, requires a detailed understanding of the distribution of risk within their portfolio, accounting for hazard (likelihood of an event occurring), vulnerability (likelihood of damage) and exposure (potential cost of damage). This process is known as loss estimation and is crucial to ensuring the financial security of insurance companies, as now required in the United Kingdom by European Union (EU) law under the Solvency II directive (Lloyds, 2015).

Economic losses resulting from an earthquake can be direct or indirect. Direct losses refer to the cost associated with repairing or replacing physical assets that have suffered structural damage and can be incurred by damage to buildings or to critical infrastructure, such as transport, energy and other utility systems. Indirect losses refer to all other measurable non-property losses. Sources of indirect losses can include: business interruption due to loss of utility supply (arising from damage to external critical infrastructure); business interruption due to break down of supply chain or unavailability of labour; provision of temporary accommodation; emergency relief costs; clean-up costs; or household spending reduction (Daniell et al., 2010). Because

building performance is more relevant to life safety, and because performance of buildings in past earthquakes has been better documented, to date there has been limited research in the field of seismic performance of infrastructure and the economic impacts of this. However, a significant proportion of losses can arise from damage to critical infrastructure and the indirect effects that disruption to their functionality has on commercial activity (Rose and Lim, 2002). Furthermore, the Sendai Framework for Disaster Risk Reduction (UNISDR, 2015) lists the reduction of economic losses and the reduction of damage to critical infrastructure and basic services as two of its seven global targets for 2030. The first priority action of the Framework is to improve understanding of all dimensions of disaster risk and if the targets are to be met, it is essential that more research is conducted into the seismic performance of critical infrastructure. Therefore, to estimate the total losses associated with an earthquake, it is necessary to understand direct and indirect losses for both buildings and infrastructure.

Loss estimation studies have been conducted by the earthquake engineering community since the 1970s (Reitherman, 1985) and are often described as seismic risk assessments. In seismic risk assessments the losses can refer to financial costs, the number of casualties, counts of property damage or some other measure of impact (Porter, 2003). Losses can be estimated for a single scenario event or account for multiple events, and they can be measured either deterministically or probabilistically. The input to a seismic risk assessment is the exposure data, which describe the physical characteristics of the assets at risk and information that will be used to measure losses e.g. asset replacement cost. The seismic risk assessment is then a three-step process including a hazard module, which simulates the earthquake event to determine a local intensity at each asset site; a vulnerability module, which estimates the level of damage experienced by each asset, based on the local hazard intensity and its physical characteristics; and a loss module, which translates the damage into an estimate of the loss metric. However, Friedman (1984) and Clark (1986) were the first to conceive the idea of developing specialist computer-based catastrophe models that could estimate losses and measure risk from natural hazard events for the insurance sector (Kozlowski and Mathewson, 1997). These have

subsequently been developed into commercial software developed by three major global vendors: Risk Management Solutions (RMS), Applied Insurance Research (AIR) and EQE by CoreLogic (Clark, 2002; Grossi et al., 2005). The specific scientific, engineering and financial details of how these models operate are generally undisclosed due to commercial sensitivity, but the frameworks applied by them follow the same structure process used by earthquake engineers for loss estimation (Grossi and TeHennepe, 2008; AIR, 2012; Botts et al., 2012).

The value of insurance is more than simply financial reimbursement. The Hyogo Framework for Action (UNISDR, 2005) prioritises the building of resilience as its objective for disaster reduction. The re-accumulation of capital from insurance pay outs acts as a catalyst for carrying out repairs and reconstruction and expedites community recovery. Tobin (1999) describes this as the 'recovery model' for building sustainable and resilient communities, whilst in the framework of Bruneau et al. (2003), insurance can build resilience through resourcefulness (providing funding) and rapidity (reducing the recovery time). Without insurance, homes and businesses may be lost, or proprietors may have to take on a significant debt burden to re-build. Tobin (1999) also describes a 'mitigation model' for building disaster resilience and insurance has a role here too, since the pricing of premiums to adequately reflect risk (i.e. high premiums in high risk areas) can be used to encourage policy holders to invest in more cost-effective measures to mitigate their risk and reduce their vulnerability (Kunreuther, 2008).

1.2 Critical infrastructure risk

The vendor catastrophe models focus primarily on losses accruing from building damage and it the process used for buildings is not directly applicable to critical infrastructure. There are two major differences in the assessment of infrastructure that require the process to be modified. The first is that infrastructure systems are large and made up of a collection of spatially distributed but connected components, which can be subjected to significantly different seismic actions over their geographical extent. The estimation of indirect losses arising from infrastructure depends not just on the performance of individual assets within the system, but also on how the

locations of damaged assets inter-relate in terms of maintaining customer supply. This affects the business operations of both the infrastructure provider itself (business interruption) and third party businesses (contingent business interruption). The importance of system performance is demonstrated by the impact on businesses after the 1994 Northridge, Los Angeles earthquake. Gordon et al. (1998) found that over a quarter of all losses affecting businesses were attributed to damage to the transportation sector, while Tierney (1997) concluded that more businesses were affected by critical infrastructure disruption than by structural damage. Losses arising from system performance issues are also not bound by the geographic extent of the physical earthquake effects. Tatano and Tsuchiya (2008) found that after the 2004 Niigata, Japan earthquake, the greatest losses did not occur in the prefecture in which the earthquake epicentre was located. Nationwide supply chains led to greater economic losses in other prefectures and so if only physical damage was accounted for, the total losses would be significantly underestimated. Therefore, for critical infrastructure, a thorough loss estimation should account for system performance, in addition to asset damage, so that business interruption losses can be adequately predicted. However, currently available commercial models do not sufficiently consider this (Rose and Huyck, 2016), which has led to the situation where loss estimates provided by catastrophe modellers in the aftermath of natural catastrophes usually exclude business interruption losses (AIR, 2016; RMS, 2016). Examples of recent natural catastrophes in which infrastructure-related business interruption losses have been significantly underestimated include the 1995 Kobe, Japan earthquake (Walker, 1997) and Hurricane Sandy, which made landfall over New York in 2012 (Ruquet, 2012).

The second difference to buildings is the existence of interdependencies between infrastructure systems, such that damage to one system may result in disruption to other systems. An example is a water supply system, which may contain pumping stations that rely on electric power to function. Another example is a natural gas delivery system, in which the main gas production plant cannot operate unless there is an available water supply for fire safety. If the two examples co-exist, then there is an indirect interdependency between the gas delivery system and the

electric power network. Furthermore, if, using the same examples, the electric power network includes a gas-fired power station, then there is a feedback relationship between the three systems. Modelling of seismic infrastructure performance can become very complex, so it is unsurprising that of the studies that do exist, many focus on a single system. However, this can lead to an optimistic prediction of performance and subsequent under-prediction of impact, so it is essential that interdependencies are identified and modelled as part of the loss estimation. Of all the steps involved in the risk assessment of critical infrastructure systems, interdependencies is a relatively recent research focus and hence it is the topic on which there is the least consensus on how it should be incorporated into analysis (Pitilakis and Kakderi, 2011). This is because the complexity of the problem has led to numerous different approaches being proposed in an attempt to simplify the problem and make interdependency modelling tractable for widespread application. The alternative approaches also reflect the fact that different sectors that are interested in interdependencies have different needs in terms of accuracy and detail; varying resources to allocate to the task; and different levels of access to model input data. The importance of interdependencies to the insurance sector does not relate just to catastrophe modelling. Many insurance brokers also provide risk management advice to clients in an attempt to mitigate losses before they occur. Where the client is an infrastructure operator, if the insurer is able to identify and quantify the magnitude of any interdependent relationships, then this information can be used to encourage the company to invest in strategies to minimise the potential impact of interdependencies. Both the catastrophe modelling and risk management considerations of interdependency are addressed in this research, and it is useful this stage to make the semantic distinction between interdependency simulation, which describes the incorporation of interdependencies into system performance analysis, and interdependency quantification, which describes the measurement of the scale of interdependency between two systems.

1.3 Research plan

The purpose of this research is to develop tools that can be used by the insurance industry to model the seismic risk to critical infrastructure systems, with specific attention being given to issues of system performance and interdependencies. Although methods for the seismic risk assessment of critical infrastructure systems do exist in the academic literature, the insurance industry has its own requirements with respect to the complexity of specific model components, which may mean that these methods are not directly applicable. Another issue is that some published methods might include model components that are location specific. Rather than relying on such models, it would be beneficial for insurers to be able to develop bespoke models as required, when empirical datasets are available. Finally, although interdependencies are considered in some existing studies, there is no consensus on how best to model them for seismic risk assessment or to quantify them for risk management.

One of the major reasons why the understanding of critical infrastructure system performance in earthquakes has lagged behind that of buildings has been the lack of available data. Operators are often unwilling to share data about their systems due to commercial sensitivity and security issues and poor damage reporting after earthquakes has made it difficult to develop and validate models. This has led to previous studies being predominantly based either on real systems, but without any empirical development or validation (e.g. Dueñas-Osorio et al., 2007; Adachi and Ellingwood, 2010), or in some cases even on theoretical systems (e.g. Wang et al., 2010; Bensi et al., 2013). There are only a very small number of studies that have validated proposed system performance models against real post-earthquake observations (Wang and O'Rourke, 2008; Javanbarg and Takada, 2009; Bonneau and O'Rourke, 2009; Wu and Dueñas-Osorio, 2013).

Consequently a major motivation for this research has been to investigate the performance of critical infrastructure with reference to a real case study for model development and validation. This has led to collaborations with the University of Canterbury in Christchurch, New Zealand and Orion New Zealand Ltd. (referred to as Orion from hereon), the electricity distribution company in Christchurch. Christchurch experienced significant damage to critical infrastructure

during the Canterbury earthquake sequence between 2010-11 and much of this damage has been recorded and documented with unusual thoroughness. Unlike in many other parts of the world, in New Zealand there is a culture of knowledge sharing in the critical infrastructure sector. This has resulted in Orion sharing exposure and observed damage data pertaining to the electric power network and the University of Canterbury providing exposure and observed damage data pertaining to the water supply system, which has been shared with them by the operator, Christchurch City Council. These two systems exhibit interdependency since the water supply system relies on electric pumps to extract groundwater from aquifers and to pump water into hilly areas of the city.

Based on the issues summarised above, this research is therefore carried out with three main objectives:

1. To review and appraise the current literature on seismic risk assessment of critical infrastructure system in order to define a 'best practice' seismic risk assessment modelling framework that includes methodologies for model development and model application. The framework will account for system performance and interdependencies and will satisfy the requirements and constraints of the insurance industry.
2. To demonstrate the application of the framework by developing an empirical model for a real case study, which will be validated and then applied for future risk projection.
3. To propose a metric and derivation method for objectively characterising and quantifying the magnitude of interdependency between critical infrastructure systems for risk management purposes.

To satisfy objective no. 2, the proposed framework model development framework (objective no. 1) will be applied to develop a catastrophe model specific to these two systems and the model is then applied to two situations: a deterministic scenario event and a probabilistic future risk assessment. The scenario event is the M_W 6.2 earthquake that struck Christchurch on 22nd February 2011, from here on referred to as the Christchurch earthquake. This deterministic

analysis mimics the regulatory stress tests and post-event rapid assessment described above. Since it is based on a real event for which damage and system performance observations exist, it also acts as a validation exercise for the developed model. However it is important to state that the value of the validation is somewhat limited because the model development stage also partly relies on empirical analysis of observed data from the Christchurch earthquake, meaning that the test and model development datasets are not independent of each other. This is unavoidable because of limitations of the datasets that has been provided. The Christchurch earthquake is the only event that has affected Christchurch for which quantitative system performances measurements are known and hence the only candidate for a validation event. For model development, observed hazard and damage data have been made available for both the Christchurch earthquake and the M_W 7.1 earthquake that struck Christchurch on 4th September 2010, from here on referred to as the Darfield earthquake, due to the location of its epicentre. However, the Darfield earthquake generated lower hazard intensities in the city and was consequently significantly less damaging. As a result the dataset from the Darfield earthquake is not sufficient on its own for model development and it is necessary to combine observations from both events to ensure that the developed model is representative of as wide a range of hazard intensities as possible.

1.4 Thesis structure

The remainder of this thesis is divided into five chapters. Chapter 2 presents a review of the existing literature relevant to the seismic risk assessment of critical infrastructure. This consists of a review of existing frameworks and algorithms for the overall process as well as a review of existing methods for the detailed science and engineering analysis required at each step. The latter includes asset characterisation for exposure data management, generation of hazard events and intensities, asset fragility and damage modelling, system performance measurement and interdependency characterisation and modelling. Although the case study in this research focuses on electric power and water supply, for the purposes of developing a more general framework, the literature review will also consider other critical infrastructure systems. Chapter

3 presents the final proposed seismic risk assessment method and model development framework. It also introduces the case study in more detail and presents the methods that will be used to validate and apply the final model and the proposed metrics and methods for quantifying interdependency. Chapter 4 describes the specific functions and methods proposed at each step to develop the case study model from the framework. These are a mixture of existing functions and methods identified in the literature or new functions to be derived specifically for the case study by empirical analysis. Existing functions and methods are to be tested for applicability against observations from the Darfield and Christchurch earthquakes and the analytical techniques used for testing are described. For new functions, the method used for derivation is described. The results from the model validation are also presented and discussed. Chapter 5 presents the results and discussion from the model application and analysis and discussion of alternative methods for quantifying interdependencies. The final conclusions from the thesis are presented in Chapter 6.

2. Literature Review

This chapter summarises the existing literature relevant to the seismic risk assessment of critical infrastructure systems. It begins in section 2.1 with an introduction to types of critical infrastructure, the features that are general to all systems, and summarises the previous studies that have conducted critical infrastructure seismic risk assessments. The subsequent sections focus on the individual components within a seismic risk assessment model. Since this thesis focuses on engineering aspects of seismic risk assessments, no discussion is provided on financial loss estimation. Section 2.2 identifies in more detail the specific aspects of each infrastructure system that must be included in an exposure database for conducting seismic risk assessments. Section 2.3 describes existing methods that can be applied for the hazard analysis, in terms of conceptual mathematical approaches, event generation, prediction of ground shaking intensity and prediction of permanent ground deformations. Section 2.4 reviews the literature on vulnerability analysis, including existing fragility functions, methods for developing new fragility functions, and the assignment of damage states. Section 2.5 provides an overview of the methods and metrics that can be used for measuring system performance in infrastructure systems. Section 2.6 explains the characteristics of interdependency and reviews existing methods for interdependency modelling and quantification. Section 2.7 discusses existing risk metrics that account for result distributions and in conclusion section 2.8 identifies the main gaps in the literature and those that will be addressed in the case study application.

2.1 Critical infrastructure systems

Critical infrastructure systems are commonly referred to within the earthquake engineering field as lifelines, reflecting their importance to earthquake relief and recovery, and the terms are used interchangeably in the literature. They can be defined as “systems or networks, which provide for the circulation of people, goods, services and information upon which health, safety, comfort and economic activity depend” (Platt, 1991). Based on a definition such as this however, there is considerable debate as to which systems qualify as critical infrastructure or lifelines. In the

absence of an agreed supranational list, different groups have assembled their own lists. Table 2.1 summarises ten studies that have sought to comprehensively define or analyse an ecosystem of critical infrastructures and the systems that they have selected.

Table 2.1 suggests that there is a core group of systems that are widely recognised as being critical infrastructure: electricity, telecommunications, highways, transportation, water supply, wastewater treatment, oil delivery, natural gas delivery, emergency services and healthcare. A characteristic common to most of these systems is that they can be represented as networks of physically inter-connected assets. Furthermore, the physically inter-linked systems can function, for at least a limited time, without human operation or participation. The exceptions are transport, emergency services and healthcare. Assets in a healthcare or emergency services system may not all be physically linked. In addition, all three of these systems require human participation, such that even if all their physical assets survived an earthquake, their performance could still be affected due to labour issues. However, this study focuses on the core group of physical networked systems, and in particular the electric power network and water supply system for the case study.

A feature of networks is that they are made up of punctual (i.e. confined to a single point in space) assets known as vertices or nodes, which are connected together by linear assets known as edges or links. Examples of vertices are electricity substations or water treatment plants, while examples of edges are water pipes or electric power transmission lines. The overall geometric form of a network is known as its topology. There are many types of network topology but the systems being considered in this study can be represented as one (or a hybrid) of two of the types: a radial network or a mesh network as shown in Figure 2.1 (Lakervi and Holmes, 1995). A radial network is characterised by a unidirectional hierarchical topology with levels of decreasing significance. The highest level contains the system sources, and at the lowest level are the system sinks, which for most systems are the customers (except for in the case of the waste water system, where the customers are the sources).

Table 2.1 – Alternative selections of lifelines presented in previous studies: 1) ATC-25 (ATC, 1991); 2) TCLEE (Schiff, 1997); 3) Kameda (2000); 4) Rinaldi et al. (2001); 5) RISK-UE (Pitilakis et al., 2006); 6) Chang et al. (2007); 7) Hazus (FEMA, 2015); 8) Porter and Sherrill (2011); 9) NIPP 2013 (DHS, 2013); 10) SYNER-G (Pitilakis et al., 2014a)

System	**Study**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Electricity	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Telecommunications	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Highways	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Transportation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Water supply	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wastewater treatment		✓	✓		✓		✓	✓	✓	✓
Oil delivery	✓	✓	✓	✓		✓	✓	✓	✓	✓
Natural gas delivery	✓	✓	✓	✓	✓		✓	✓	✓	✓
Building services						✓				
Business				✓		✓			✓	
Emergency services	✓	✓		✓		✓		✓	✓	✓
Financial systems				✓		✓			✓	
Food supply				✓		✓			✓	
Government				✓		✓			✓	
Health care	✓	✓		✓		✓		✓	✓	✓
Solid waste			✓							
Education				✓						
Space				✓						
Commodities				✓						
Nuclear power									✓	
Dams		✓						✓	✓	
Chemical facilities									✓	
Manufacturing									✓	
Defence bases									✓	

There can be any number of levels in between and usually the number of edges and vertices in each level increases from source to sink as the network branches out, hence radial networks are often described as ‘tree-like’. A simple case is the electric power network shown in Figure 2.1, which begins at ‘level 1’ with a single source node, the generation plant. This feeds two substations in ‘level 2’, each of which feeds a larger number of sink nodes. Every node (except source nodes) should be connected to at least one node in the level above and there should be a path to every sink node from at least one of the source nodes. Usually in radial networks there is no interconnection between nodes at the same level. In general, although they are efficient in their asset requirements, they often have very low levels of redundancy so that they run a high risk of system failure (Leelardcharoen, 2011). Mesh networks are often described as ‘grid-like’ and although they can be characterised in terms of a hierarchy for management purposes, in operational terms any node at any level can act as a source or a sink. Mesh networks, as shown in Figure 2.1 are predominantly bi-directional (although they may contain some unidirectional links), meaning that there are many interconnections between nodes at all levels, including between nodes at the same level. Mesh networks require more physical assets but the system has greater redundancy, reducing risk of system failure.

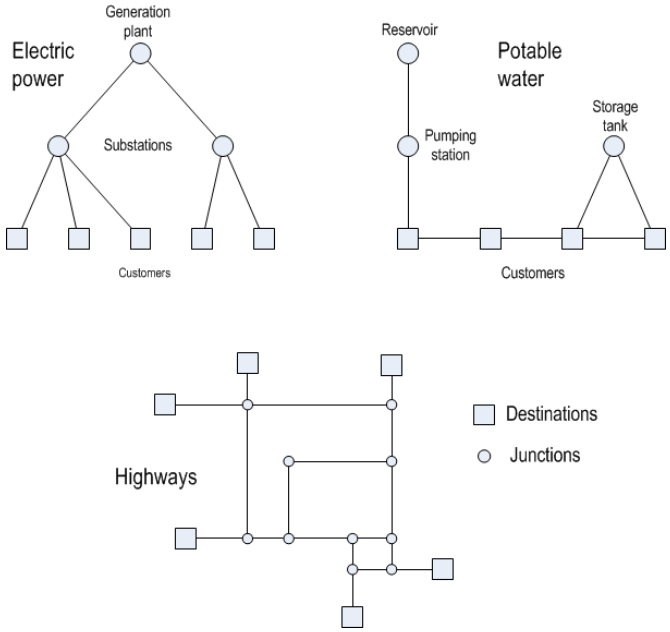

Figure 2.1 – Examples of network types including hierarchical (electric power and potable water supply) and mesh (highways)

The inter-connected nature of infrastructure systems is important for loss estimation. For a portfolio of buildings, the total loss from an earthquake can be determined simply by aggregating the expected losses from each individual building. For infrastructure systems however, the losses depend on the overall performance of the system, i.e. the ability of the system to fulfil its customer demand, which depends on both damage to individual assets and on how the assets inter-relate. Therefore the method for seismic risk assessment of buildings must be adapted for application to critical infrastructure systems. Table 2.2 lists previous academic and institutional studies on the seismic risk of core critical infrastructure systems including both specific case studies and more general methodological frameworks. A basic description of the methodologies applied in each study is provided in the adjacent columns. The second column lists the systems considered in the analysis where: EP = electric power; WS = water supply; WW = wastewater collection; NG = natural gas; OD = oil delivery; T = telecommunications; and H = highways. The third column refers to whether the loss metric is calculated for a single event or as a composite metric for multiple events. The fourth column lists the general types of loss estimates considered in the analysis, where: AR = asset restoration time; BH = building habitability index; C = number of casualties; DC = damage count; DEC = direct economic costs; IEC = indirect economic costs; SD = system downtime; SP = system performance. Note that since asset damage counts (or some probabilistic equivalent) are a necessary initial step to calculate other metrics, the designation ‘DC’ is only listed where damage counts are the only loss metric considered. The fifth column refers to whether the loss metrics are presented deterministically (D) or probabilistically (P). The final column describes whether the analysis considers interdependencies between systems.

Table 2.2 – Summary of previous studies conducted on seismic risk assessment of critical infrastructure systems

Study	**System(s)**	**Event**	**Loss metric(s)**	**Analysis**	**Inter-dependencies**
ATC-13 (ATC, 1985)	EP, WS, WW, NG, H, T	Single	DEC, IEC, SP, SD	D	No
ATC-25 (ATC, 1991)	EP, WS, WW, NG, OD, H	Single	DEC, IEC, SP, SD	D	No
Shinozuka and Tanaka (1996)	EP, WS	Single	SP	D	Yes
Vanzi (1996)	EP	Multi	SP	D	No
Hwang et al. (1998)	WS	Single	SP	D	No
Kawakami (2000)	H	Single	SP	P	No
American Lifelines Alliance (2001)	WS	Single	DC	D	No
Chang et al. (2002)	WS	Single	SP, SD, DEC, IEC	D	No
American Lifelines Alliance (2005)	EP	Multi	DEC, IEC, SP, SD, C	P	No
RISK-UE (Pitilakis et al., 2006)	EP, WS, WW, NG, H, T	Single / Multi	DEC, IEC, SP, AR	D, P	Yes
Wang and O'Rourke (2006)	EP, WS	Single / Multi	SP	D, P	Yes
Chang and Song (2007)	NG	Single	SP	P	No
Shinozuka et al. (2007)	EP	Multi	SP	P	No
Shiraki et al. (2007)	H	Multi	SP	P	No
Adachi and Ellingwood (2008)	EP, WS	Single	SP	P	Yes
Rasulo et al. (2008)	EP, WS, H	Single	SP	D	No
Azevedo et al. (2009)	EP, WS, NG, OD, H	Single	DC	D	No
Hazus-MH (FEMA, 2015)	EP, WS, WW, NG, OD, H, T	Single / Multi	DEC, IEC, AR, SP	D, P	No
Bonneau and O'Rourke (2009)	EP, WS	Single	SP, SD	D	Yes
Guikema and Gardoni (2009)	H	Single	SP	P	No

Study	System(s)	Event	Loss metric(s)	Analysis	Inter-dependencies
Javanbarg and Takada (2009)	WS	Single	SP	P	No
Reed et al. (2009)	EP, T	Single	SP, SD	D	Yes
Tang and Wen (2009)	WS	Multi	DEC, C	D	No
Jayaram and Baker (2010)	H	Multi	SP	P	No
Kim et al. (2009)	EP, WS	Multi	SP	P	Yes
Wang et al. (2010)	WS	Single	SP	P	No
Hernandez-Fajardo and Dueñas-Osorio (2011)	EP, WS	Multi	SP	D, P	Yes
Leelardcharoen (2011)	EP, T	Single	SP	P	Yes
Poljansek et al. (2012)	EP, NG	Multi	SP	P	Yes
Wu and Dueñas-Osorio (2013)	EP, WS	Multi	SP	P	Yes
Cavalieri et al. (2014a)	EP	Multi	SP	D	No
SYNER-G (Pitilakis et al., 2014a)	EP, WS, WW, NG, OD, H	Multi	SP, C, BH	P	Yes
Esposito et al. (2015)	NG	Multi	SP	P	No

Of particular interest are the five studies in Table 2.2 that propose a methodological framework for a wide range of physical systems, which could potentially provide a basis for the framework proposed here. ATC-13 (ATC, 1985) was developed as a set of guidelines for seismic damage evaluation in California and was the first study to include a proposed methodology for seismic risk assessment of critical infrastructure. This was later developed into ATC-25 (ATC, 1991) which focused specifically on critical infrastructure but for the whole of the United States. Eventually these guidelines evolved into the Hazus-MH (abbreviated to Hazus from here on) multi-hazard risk assessment software (NIBS, 1999; FEMA, 2003; FEMA, 2015). One weakness of Hazus is in the estimation of indirect losses, since it does not account for interdependencies between systems and business interruption losses are estimated at sectorial level rather than for individual businesses. It also does not consider reduction in business

functionality due to infrastructure damage (contingent business interruption), only business interruption due to building damage. The impact of infrastructure on indirect losses is only considered from the perspective of more rapid infrastructure repair being a catalyst for more rapid building repair. The RISK-UE study (Pitilakis et al., 2006) was initially developed with the intention of creating a European equivalent to Hazus. The final output from the project was a methodological framework, which although thorough in terms of exposure and fragility information, is relatively weak on methods for calculating loss metrics and interdependencies. This was followed by the SYNER-G study (Pitilakis et al., 2014a), which set out with the same aims as RISK-UE but with the benefit of the prior work undertaken, was able to focus more strongly on loss metrics and does include analysis of unidirectional interdependencies.

Although each of the studies in Table 2.2 varies in terms of structure and methodology, the commonality between all of them is that they follow the same basic process as the seismic risk assessment of buildings described in Chapter 1, with hazard, fragility and loss calculations being applied to an exposure database. The following sections provide a detailed technical description and critique of the existing methods used for each of these stages, both in the studies listed in Table 2.2 and other studies relevant to each stage.

2.2 Exposure

2.2.1 System classification

The purpose of an exposure database is to identify the assets at risk and the physical attributes of each asset that need to be known in order to determine the vulnerability of the asset to earthquakes. In addition to physical attributes, infrastructure assets also possess one or more systemic attributes, i.e. the function that the asset performs in the system. Therefore four pieces of information are needed to describe assets in an infrastructure system: its location, its function, its physical attributes; and the adjacent assets to which it is connected. Whilst the location and connections of an asset are straightforward to determine, the function and physical attributes require specialist knowledge of the system to which it belongs. One objective of the

framework to be proposed in this study is to provide an insurer with all the information that could help describe and characterise an infrastructure system for seismic risk assessment. This means that it is important to list all of the assets that could potentially exist in a system, together with the important descriptive attributes related to that asset, that would ideally be needed for the assessment of seismic losses. Appendix B includes tables detailing the infrastructure classification schemes adopted by the studies listed in Table 2.2 and these are summarised in the following sub-sections.

2.2.2 Electric power

The supply of electrical power can be divided into a three-stage process involving generation, transmission and distribution (Lakervi and Holmes, 1995). Generation is the conversion to electrical energy from alternative energy sources; transmission is the process of transferring large amounts of electrical energy from generation areas to sites of demand; and distribution is the process of providing electricity to the customer within the demand centres. Electric power leaves generation plants at high voltages, which are stepped down along their route by transformers, in order to provide power with a usable voltage level. The reason for transmitting electricity at high voltages from generation plants is that this allows transmission at a lower current, which means that thinner cables can be used, thus reducing resistance and power losses (Cavalieri et al., 2014b). Some studies focus on regional or national assessments that consider only the transmission network and treat the demand centres as end nodes (e.g. Vanzi, 1996; Pitilakis et al., 2014a). Instead urban studies are more likely to focus on local distribution networks and potentially exclude the transmission network altogether if generation plants are far from the city, such that the probability of both the generation plant and the local distribution network being damaged at the same time is very low. The majority of studies on the electric power network focus on three basic system functions: generation plants, substations and conduits.

Generation plants are the sites where electric power is produced such as a fossil fuel power station or a hydro-electric power plant. ATC-13 (ATC, 1985) treats generation plants as buildings so that the only physical attribute used to distinguish between different plants is the building construction class. Later studies (e.g. Pitilakis et al., 2014a; FEMA, 2015) treat generation plants as separate entities to buildings and most of these identify capacity and whether or not the plant has been seismically designed as the key physical attributes to inform vulnerability analysis. Not all studies include generation plants in their classification scheme. In the case of Reed et al. (2009) this is because of a deliberate focus on lines, but in the other cases this is because the generation plants are located far from the actual case study area and therefore the generation plant and the case study system are not considered to be simultaneously vulnerable.

Substations are intermediate nodes in an electric power network and can perform one of two functions (Cavalieri et al., 2014b). They can be distribution substations, which usually receive power from a single high capacity conduit and re-distribute the power to a larger number of lower capacity conduits. Distribution substations move power from one hierarchy level to the next. Transformation substations perform the same distribution function and additionally convert power from high voltage to low voltage. As with generation plants, ATC-13 (ATC, 1985) treats substations as buildings, differentiated by their construction class. Later studies (e.g. Pitilakis et al., 2014a; FEMA, 2015) treat substations separately and most of these identify voltage and whether or not the station has been seismically designed as the key physical attributes. Substations can exist at different levels in a network hierarchy and there are different approaches to how this is treated. Some studies identify substations at different levels as distinct functions, whilst other studies identify different levels as distinct attributes of a single substation function. Cavalieri et al. (2014a) and Pitilakis et al. (2014a) treat substations in a completely different way. They do not consider a substation to be a single entity and instead treat it as an inter-connected collection of small micro-components, each with its own function and set of attributes to be analysed individually. This allows for more precise vulnerability analysis and

more sophisticated system performance analysis, making it very useful for infrastructure operators themselves to inform their own risk assessments. However, it is questionable whether this method has value for the insurance sector given the requirements for data describing the internal layouts of each substation in a network.

Conduits are the linear assets that transmit or distribute electricity between nodes and can be either buried or above ground (overhead), supported by towers or poles. However, there is a lack of consensus on terminology in the literature. Some studies use the term ‘lines’ as a general term to refer to both buried and overhead assets (e.g. Pitilakis et al., 2006), whist other studies (e.g. Adachi and Ellingwood, 2008) use the term ‘lines’ to refer specifically to overhead assets, with the term ‘cables’ used to describe buried assets. For avoidance of doubt, in this study the latter classification is adopted, so that overhead linear assets are ‘lines’, buried linear assets are ‘cables’ and ‘conduits’ is used as a general term for all types of linear asset. Two studies make a functional distinction between conduits in the transmission network and conduits in the distribution network (ATC, 1991; ALA, 2005) and only three studies identify any physical attributes associated with conduits, in this case whether or not they have been seismically designed (Pitilakis et al., 2006; Azevedo et al., 2009, FEMA, 2015).

Four studies consider an asset known as a gate station, although it is notable that all of these use the same case study of the electric power network in Memphis. It is apparent that these studies also do not include generation plants in their assessment because of distance and the name ‘gate station’ is simply a label given to the substations at the entry points to the city network from the generation plants. Therefore in physical and functional terms, they are same as substations. ALA (2005) consider office buildings and control equipment in their analysis and Leelardcharoen (2011) also considers a control centre. Since electricity cannot be stored, control centres play a critical role in networks in balancing supply with demand to avoid wastage. However the failure of a control centre or control equipment would not cut off supply to a city and so has negligible impact on system performance and indirect losses.

2.2.3 Water supply

A water supply system can be divided into two parts, the transmission system and the distribution system (Kakderi and Argyroudis, 2014). In the transmission system, raw water is taken from its source to a treatment plant. The distribution system delivers treated water to customers and is often divided into a hierarchy of the mains system and submains system. Other elements of a water supply system that feature regularly amongst the studies in Table 2.2 include pumping stations, storage facilities (e.g. reservoirs and tanks), and conduits to carry water such as pipes, tunnels and canals.

Water sources include wells, springs, lakes, rivers and impounding reservoirs. However, from the perspective of post-disaster system performance, it is the engineered assets such as wells and reservoirs that are of primary interest. The majority of studies that consider sources refer specifically to wells. Only ATC-25 (ATC, 1991) and Wang et al. (2010) consider reservoirs, although RISK-UE (Pitilakis et al., 2006) and Javanbarg and Takada (2009) consider the generic entity 'sources'. One of the reasons why source reservoirs may be so uncommon is that although they function as a water source, in physical terms they are more similar to a storage facility. For damage estimation, the physical attribute of an asset is more important than the functional attribute, so many studies treat reservoirs as a type of storage facility, irrespective of their systemic function. The RISK-UE study (Pitilakis et al., 2006) considers five attributes for wells: capacity, whether it requires a pump, whether it has back-up power for the pump (if applicable), the building construction class and whether it is seismically designed. The SYNER-G study (Pitilakis et al., 2014a) also considers the issue of seismic design as an attribute but is the only other study to consider an attribute for wells.

Treatment plants are complex structures in which a number of physical and chemical processes take place to improve the quality of the water. Whether treatment is necessary depends on how the water is obtained and it is not uncommon in places where water is extracted from deep aquifers (e.g. in Christchurch, New Zealand) for water to be delivered to customers untreated.

As a result, some of the water system studies in Table 2.2 do not include treatment plants in their classification schemes. ATC-13 (ATC, 1985) considers treatment plants as buildings and so building construction class is used to distinguish between them. Later studies treat treatment plants (e.g. Pitilakis et al., 2014a; FEMA, 2015) as separate entities to buildings and some of these identify capacity and whether or not the plant has been seismically designed as the key physical attributes.

Pumping stations are used to increase water pressure either because it is extracted at low pressure or to distribute water into elevated areas where pressure head would otherwise be lost. Whether pumping stations are necessary depends on the water source and the topography of the study area. However, nearly all the studies in Table 2.2 include pumping stations in their classification scheme. The exceptions are Rasulo et al. (2008) who focus on pipes and, without explanation, ALA (2001). ATC-13 (ATC, 1985) considers pumping stations as buildings and so building construction class is used to distinguish between them. Later studies treat pumping stations (e.g. Pitilakis et al., 2014a; FEMA, 2015) as separate entities to buildings and most of these identify capacity and whether or not the station has been seismically designed as the key physical attributes. However the RISK-UE study (Pitilakis et al., 2006) also considers the building class and whether or not the station has back-up power as attributes. The latter particularly is important if interdependencies with electric power systems are being considered.

Storage facilities include covered tanks and open reservoirs, which can in turn be engineered either artificially or from the natural landscape. Although not all systems will contain storage facilities, all the water system studies in Table 2.2 consider them except Chang et al. (2002), Rasulo et al. (2008) and Tang and Wen (2009). There are numerous attributes associated with storage tanks that are considered by these studies including: elevation, material foundation, whether the tank is anchored, the type of roof, storage capacity and whether it is seismically designed. It is questionable however whether all of this information would be available to an insurer for all tanks in a system and so it may be more realistic for an insurance-focused framework to focus on the most important attributes. In this regard, elevation, material and the

anchorage of the tank are the three attributes that feature most commonly in existing studies and so could be interpreted as being the most important. Only ATC-25 (ATC, 1991) and Wang et al. (2010) explicitly consider storage reservoirs as distinct entities from storage tanks. Neither of these considers any attributes for reservoirs.

Water conduits can take the form of pipes, tunnels, canals or aqueducts. Pipes are the most common and feature in all the water system studies in Table 2.2. Material, and to a lesser extent, diameter are the most common attributes considered for pipes. Other attributes considered include: type of joint (rigid or flexible), soil conditions, elevation (buried or above ground) and whether the pipe has corrosion protection. Canals and tunnels feature in three studies: ALA (2001), RISK-UE (Pitilakis et al., 2006) and SYNER-G (Pitilakis et al., 2014a). For tunnels, the key attributes are the construction type and the local ground conditions. For canals, the RISK-UE and SYNER-G studies focus on the construction type and whether the canals are reinforced. However, the ALA study considers maximum flow capacity, complex geometric parameters and the locations of in-line components such as valves and siphons. Aqueducts only feature in the ATC-25 (ATC, 1991) and ALA studies. In the ALA study, elevation and the locations of pumps and appurtenances are considered to be important attributes.

2.2.4 Waste water

A waste water collection system operates in four parts: the collection of waste water, its transmission to treatment plants, its treatment and finally the disposal of the treated water (Kakderi and Argyroudis. 2014). Only five studies consider waste water systems: ATC-13 (ATC, 1985), ATC-25 (ATC,1991), RISK-UE (Pitilakis et al., 2006), Hazus (FEMA, 2015) and SYNER-G (Pitilakis et al., 2014a). The main components of a waste water system included in all these studies are the conduits for the collection, transmission and disposal of the waste water, treatment plants and pumping stations (also known as lift stations) which are needed to transport water over topographical barriers. These components are treated in the same way as the equivalent components in the water supply system.

2.2.5 Natural gas

A natural gas delivery system can be considered to include: the production and processing of natural gas followed by its transmission and distribution (Gehl et al., 2014). Natural gas is extracted from underground wells and must be purified before it is distributed, hence the need for processing. The gas must then be pressurised to provide the energy for transmission to demand centres. Pressure reduction is then required to distribute the gas to customers at an appropriate pressure. Transmission and distribution occurs in pipelines, which are usually underground and known as buried pipelines. Levels of gas transmission vary according to seasonal demand, with peak season in winter due to increased heating requirements and lower demand in the summer (Esposito, 2011). To account for this variation, storage facilities are used with excess gas produced in summer being stored for use in winter. Storage facilities tend to be underground, such as depleted gas fields, depleted aquifers or salt caverns. However natural gas can also be liquefied and stored in above ground tanks. Nine of the studies in Table 2.2 consider natural gas systems. The key components included in all studies are compressors stations and pipelines. Four of the studies also include storage tanks, but only one study, RISK-UE (Pitilakis et al., 2006), includes the primary processing plant, perhaps because this activity is likely to take place away from urban areas and has therefore been neglected by authors studying urban systems. Esposito et al. (2015) also consider regulators but this is a type of compressor station that exists in the case study area (Italy) at low pressures closer to the customers. The attributes for pipelines and tanks are same as for their equivalents in the water supply systems. For compressor stations, the only attribute considered by most studies is whether it is seismically designed. The RISK-UE study also considers seismic design as an attribute for processing plants.

2.2.6 Oil delivery

Oil delivery systems can be considered to be analogous to natural gas delivery systems with a production and processing phase followed by a transmission and distribution phase (Gehl et al.,

2014). Only four studies consider oil delivery systems: ATC-25 (ATC, 1991), Azevedo et al. (2009), Hazus (FEMA, 2015) and SYNER-G (Pitilakis et al., 2014a). All of these studies consider pipelines and tank farms in their classification, with the same attributes as their equivalents in the natural gas system. All the studies except ATC-25 consider pumping stations whose attribute is whether it is seismically designed. All studies except SYNER-G consider refineries, whose attributes are capacity and whether it is seismically designed. In addition, ATC-25 separately considers oil fields with no identifying attribute. The exclusion of this asset from other studies is not specifically explained but may be because oil fields only exist in a small number of places around the world and so are not relevant to the majority of case studies. Even where they do exist, their engineered assets could be represented by a refinery or pumping station (or combination of), depending on the specific activities that take place there. Ports are often a critical part of the supply chain for oil, however these are usually considered separately in studies.

2.2.7 Telecommunications

Wired telecommunications systems can be divided into local networks, which cover a small geographic area, and toll networks which link local networks together (Leelardcharoen, 2011). Local networks interface with toll networks at offices called access tandems. Within local networks, central offices provide links between customers. The last central office in the network is directly connected to the customer via telephone lines or loops, and is called an end office. End offices are not directly connected to each other, but are connected via higher-level central offices. The connections between central offices and end offices are known as trunks. Within mobile communications, mobile phones connect to each other via base stations (Leelardcharoen, 2011). Groups of base transceiver stations are connected to a base station controller all of which forms a base station subsystem. Many base station subsystems connect to a gateway switching centre, which then connects to the wired system. Because of the number of transmitters required in a mobile network, many are mounted to buildings, which means the seismic risk is related to the building rather than the transmitter itself. Five studies consider telecommunications systems:

ATC-13 (ATC, 1985), RISK-UE (Pitilakis et al., 2006), Hazus (FEMA, 2015), Reed et al. (2009) and Leelardcharoen (2011). The main components considered by these studies are overhead lines, buried cables and offices. ATC-13 does not consider lines or cables, but does consider transmission towers. The RISK-UE study considers material, burial depth and installation method as attributes for cables. For offices, seismic design (Hazus), anchorage of internal equipment (RISK-UE) and building construction class (ATC-13) are potential attributes.

2.2.8 Highways

The links in a highways system are the paved roads, which can either exist as roads by themselves or along bridges and tunnels. Ordinary roads are considered in most studies and the most common attribute is simply the level of the road in the hierarchy. However, the RISK-UE study (Pitilakis et al., 2006) also considers the geometry and topography of the road and whether there are buildings alongside the road. Bridges are also considered in most studies and possess a long list of potential attributes including: road hierarchy level, construction type, pier type, connection type, bearing type, span length, age, number of spans, skew, height, foundation type, maximum span length, number of column alignments, presence of adjacent cross-over (or cross-under) bridges, width and whether it is seismically designed. In practice however, it is likely that much of this information would not be available to an insurer modelling an entire system. Tunnels are also considered by most studies and its main attributes are construction method and ground conditions. Unlike other systems, at most nodes (junctions) there are no specific physical assets, although traffic signals will occur at some junctions. However, although traffic signals may cause a highways system to operate less efficiently, they would not cause a significant performance reduction in the highways system and so they are generally not considered in analyses. Other potentially significant elements of a highways system include embankments, trenches and slopes. However, this study is focused on infrastructure in an urban setting and so examples of this are likely to be rare and are not considered further.

2.2.9 Concluding remarks

This section has described in general terms, the main assets that exist in all of the critical infrastructure systems and the key attributes of each asset are included in Appendix B. The specific assets and corresponding attributes associated with the systems being analysed in the case study for this research are presented in more detail in Chapter 3.

2.3 Hazard

2.3.1 Conceptual approaches

Historically there have been two major conceptual approaches to seismic hazard assessment: deterministic seismic hazard assessment (DSHA) and probabilistic seismic hazard assessment (PSHA) (Reiter, 1990). In DSHA, a single reference earthquake with specific characteristics is used. It is assumed that the probability of occurrence of that earthquake within the design life of an asset is one (Bommer, 2002). Any event can be chosen to be the reference earthquake and possibilities might include the maximum credible earthquake or an event with characteristics matching a significant historic event in the region. Both of these options rely on expert judgement. The output of DSHA is an estimate of earthquake intensity (ground shaking and/or ground deformation) across the region of interest, representing the effects of a specifically defined earthquake. As described in Chapter 1, analysis of single events is an integral part of the activities of an insurer, for regulatory requirements and for rapid post-earthquake assessments. However, analysis of only a single event means that the uncertainty relating to which events may actually occur over the life an asset are not taken into account, and therefore DSHA does not provide a true representation of risk. Uncertainties in the evaluation of seismic hazard include how many earthquakes might occur, their epicentral locations, their magnitudes and other characteristics of the strong ground shaking that may affect the likelihood of assets being damaged. Whilst it is possible at some locations that only one fault exists that is likely to generate an earthquake, at most locations an asset will be within range of multiple faults. Knowing which fault has generated an earthquake is important since the magnitude is related to

the geometric properties of the fault and both magnitude and source-site distance affect the intensity of ground motion at the site of interest and thus have a direct effect on asset damage probability.

In PSHA, developed originally by Cornell (1968), the risk associated with multiple potential earthquake sources is determined by first understanding the average activity rates of each source and then estimating the ground motion intensities that are likely to occur across a region with a specified probability of exceedance within a specified time period (known as a return period if the probability is 1). This method has become the most widely used for seismic risk assessment (Bommer, 2002). Since it accounts for the uncertainties in earthquake characteristics, the merits of PSHA for representing overall risk are clear. However, when conducting PSHA for a region rather than for a single site, the ground motion intensities generated by the analysis are a statistical artefact. This means that the ground motion intensities calculated for two independent sites in the region may have the same return period but are not necessarily caused by the same earthquake event, particularly if the two sites are distant from each other. The ground motion intensities presented in a PSHA map are a composite statistical representation of all possible earthquakes, but do not realistically replicate what might be caused by a single event. For spatially distributed infrastructure systems this is problematic. Suppose that there is a system which fails only if two specific components, which are geographically distant, are both completely damaged at the same time. PSHA tells us the ground motion intensities that will occur at those two locations for a specified return period but not the simultaneous ground motion intensities, so system performance, and hence indirect losses, cannot be calculated. Because PSHA ground motion intensities are usually calculated for long return periods and account for geographically spread sources, they are conservative compared to a realistic single event ground motion intensity map which would exhibit attenuation of ground motion intensity for sites further away from the source. This would increase the likelihood of predicting asset failures and potentially result in a more conservative estimate of indirect losses.

An alternative to using either DSHA or PSHA is to use a hybrid approach that combines the optimal aspects of both (Bommer, 2002). In the hybrid approach, realistic specific events are modelled to ensure that simultaneous intensities are estimated, but to account for uncertainties, this is done for a set of events rather than for a single event. The simulated set of events should be probabilistically representative of all potential earthquakes in a study area (a stochastic event set). This means that the distributions associated with the uncertain variables (e.g. location, magnitude) should be replicated in the catalogue and this requires that the catalogue is sufficiently large. Final risk measures relating to losses can be determined by statistical analysis of results from all events in the set.

Disaggregation of seismic hazard is a method that can be used to analyse scenario events within a PSHA framework (Bazzurro and Cornell, 1999). For a ground motion with specified return period, the disaggregation method identifies which events from the hazard source model contribute to that level of ground motion and the proportion of the contribution. However, whilst this method works well with single-site analyses, its applicability to spatially distributed assets is questionable, since the events that produce the specified return period ground motion intensity for one asset will not necessarily produce ground motion intensities of the same return period for other assets.

Both RMS (Grossi and TeHennepe, 2008) and AIR (2012) software use stochastic event sets as the basis for their hazard modules, even for risk assessment of buildings. The justification for this is that insurers are generally interested in the risk faced by a portfolio of buildings, rather than by a single building, and in particular the likelihood and potential implications of multiple buildings being damaged simultaneously. As with infrastructure systems, this requires realistic event scenarios. Given that hybrid methods are the most appropriate for infrastructure systems and that existing catastrophe models already adopt this method anyway, the following sub-sections focus on the components of hybrid methods with specific consideration of their application to critical infrastructure seismic assessment.

2.3.2 Stochastic event set

Since the stochastic earthquake set is intended to represent the collective risk of all potential sources in a study area, it is necessary to define the potential sources and determine the activity rates for each source. This requires knowledge of past earthquakes in the study area. Such information can come from two sources: historic catalogues (anecdotal with magnitude estimated) and instrumental catalogues (earthquakes recorded by seismometers with magnitude calculated) (Coburn and Spence, 2002). Instrumental catalogues are better for low magnitude earthquakes, which are commonly missed by anecdotal records. However instrumental seismology only began becoming widespread in the mid-to-late 20th Century, so this is of limited use for larger magnitude earthquakes with long return periods. Therefore to create a complete catalogue across the magnitude range, records from both are required.

In some study areas it may be possible to identify the specific source (either point source or fault) associated with an earthquake. However, some earthquakes may have occurred at sites where there is no identifiable source. These events contribute to what is known as distributed or background seismicity and the most common way of dealing with them when compiling a catalogue is to divide the study area into areal source zones (Coburn and Spence, 2002). The boundaries between source zones are subjective, but judgement can be informed by the observation of earthquake clusters and tectonic and geological information. A defining feature of a source zone is that there is an equal probability of an earthquake occurring at any point within the zone.

When studying the activity rates of earthquake sources, Gutenberg and Richter (1944) found that the distribution of magnitudes generally followed a linear relationship as shown in Eq. 2.1, where N_{GR} is the annual frequency of events with magnitude greater than or equal to M and a_{GR} and b_{GR} are coefficients to be determined by model fitting. This is known as the Gutenberg-Richter relation and the minimum information required in a catalogue to determine it, is the date, location and magnitude of each event. Another method for describing the activity of a

specific point or fault source is the characteristic earthquake model. Schwartz and Coppersmith (1984) postulated that some sources have a tendency to repeatedly generate earthquakes of, or very close to, a defined characteristic magnitude.

$$\log_{10} N_{GR} = a_{GR} - b_{GR} M \tag{2.1}$$

Grossi and Windeler (2005) describe two ways of implementing a stochastic event set in catastrophe modelling. The first is a logic tree approach, which includes every possible earthquake from every source in the event set. In the final loss analysis, the results generated from each event are weighted according to the annual probability of each event occurring, as determined from the Gutenberg-Richter relation derived for the corresponding source. The second approach is a Monte Carlo simulation, which can be implemented as either a specified number of events or as a set of events within a specified number of years. The source of each event can be assigned by randomly sampling from all sources, with weights equivalent to the overall annual frequency associated with each source. Similarly the magnitude of each event can be assigned by randomly sampling from the distribution represented by Gutenberg-Richter relation for the assigned source. If the simulation is of years rather than events, then the time between events in the set needs to be modelled. It is commonly assumed that occurrences of earthquakes are independent of one another and that their recurrence pattern follows a Poisson process with a mean inter-arrival time determined from the catalogue (e.g. Cornell, 1968; Coburn and Spence, 2002). The time between events can then be modelled by randomly sampling from an exponential distribution with the mean inter-arrival time as its rate parameter.

2.3.3 Ground shaking intensity

The expected damage caused to physical structures after an earthquake is related to the severity of ground shaking caused by seismic waves. As seismic waves spread out from a source there is a reduction in amplitude and subsequent loss of energy that causes the ground shaking intensity to attenuate with distance (Bolt, 2006). The purpose of ground shaking intensity measures (IMs) is to quantify the severity of ground shaking at sites of interest so that they can be used to relate

loss or damage to hazard. Ground shaking intensity can be measured as either felt intensity or instrumental intensity. Felt intensity (also referred to as macroseismic intensity) is a subjective measurement based on observed damage to structures and human perception. Although felt IMs can be used for seismic risk assessment, their subjectivity leads to significant uncertainty. For engineering purposes, it is more desirable to use instrumental IMs that objectively record physical characteristics of the seismic waves that cause ground shaking. Typically, strong motions seismographs record the acceleration of ground shaking over the duration of an earthquake. The resulting accelerogram, which is a composite of all the seismic waves produced in an event, is the basis of instrumental IMs. Three types of IMs can be interpreted from a accelerogram: amplitude parameters, duration parameters, and frequency-content parameters.

Peak ground acceleration (PGA) and peak ground velocity (PGV) are two amplitude parameters that are commonly used in seismic risk assessments and fragility studies. PGA can be measured directly as the maximum amplitude of the accelerogram. PGV can be measured as the maximum amplitude after integration of the accelerogram. Peak ground displacement (PGD) can be measured after double integrating the accelerogram, but this IM is rarely used in engineering studies. Duration parameters are important since structures can suffer significant damage even when amplitude of shaking is moderate, if the duration of shaking is long. An example of a duration parameter is bracketed duration, which measures the time between the first and last peaks of 0.05 g (0.49 m/s^2) on an accelerogram. While amplitude and duration parameters measure ground shaking as a single value that represents all seismic waves in an earthquake, frequency-content parameters individually measure the shaking effect of the different wave frequencies in the earthquake. Frequency-content parameters can be represented graphically as response spectra, which are envelopes of the peak response at each frequency-level. The measured response can be in terms acceleration, velocity or displacement and are referred to as spectral acceleration (SA), spectral velocity (SV) and spectral displacement (SD).

For seismic risk assessments it is necessary to predict ground shaking severity at asset sites in terms of one of these IMs. Ground motion prediction equations (GMPEs) are models that

predict the ground shaking IM at a site based on magnitude, source-site distance and additional relevant characteristics (e.g. Youngs et al., 1997; Campbell, 2003; Akkar and Bommer, 2010). e.g. depth, soil conditions, fault type. The typical mathematical form of GMPEs is shown in Eq. 2.2 (Bradley, 2013), where $\boldsymbol{\theta}_{\text{GMPE}}$ is a vector of predictor variables including magnitude, source-site distance and other relevant characteristics (e.g. depth, soil conditions, fault type etc.), $\boldsymbol{\beta}_{\text{GMPE}}$ is a vector of the derived model coefficients, η is the inter-event residual and ε is the intra-event residual.

$$\ln \text{IM} = f\left(\boldsymbol{\theta}_{\text{GMPE}} \mid \boldsymbol{\beta}_{\text{GMPE}}\right) + \eta + \varepsilon \tag{2.2}$$

The two residual terms represent the uncertainty associated with the empirical derivation of the GMPE. Inter-event variability arises from separate earthquakes with identical characteristics causing different ground shaking intensities at comparable distances. For a single earthquake, the inter-event variability parameter is constant at all sites. Intra-event variability arises within a single event and represents variation of ground shaking intensity at sites that are equidistant from the earthquake source. The residual terms are random variables, where η has zero mean and standard deviation τ, and ε has zero mean and standard deviation σ. Most ground motion prediction equations described by Douglas (2015) define specific values of σ and τ. In the seismic risk assessment studies listed in Table 2.2, the most commonly adopted ground shaking IM is PGA, although PGV is sometimes also adopted since it is a good predictor of damage in pipes (Pineda-Parros and Najafi, 2010). The likely reason for the prevalence of PGA is that it is the IM for which the most GMPEs exist and the majority of fragility functions for critical infrastructure assets predict damage probabilities in terms of PGA. A comprehensive list of all GMPEs derived between 1964 and 2015 has been compiled by Douglas (2015) and includes details of 384 GMPEs for PGA and a further 246 for SA. A GMPE is categorised by the tectonic environment for which it has been developed. Delavaud et al. (2012) identify six such tectonic environments: stable continental, oceanic crust, active shallow crustal, subduction zones, deep-focus non-subduction earthquakes and active volcanoes. When selecting a GMPE,

if no local equation is available, Delavaud et al. (2012) advise selecting one developed for the same tectonic context.

It is common in studies on spatially distributed systems to either assume that all sites are perfectly correlated (i.e. the residual terms are constant across all sites) or that all sites are completely independent (i.e. the residual terms are generated randomly), but this can lead to overestimation or underestimation of losses respectively (Weatherill et al., 2014). Moreover neither assumption is realistic. For a single event, some spatial correlation of intensities will occur at sites which are close together due to correlation of intra-event residuals, as demonstrated by Boore et al. (2003). However, it is unlikely that ground conditions and seismic wave characteristics would be sufficiently uniform for perfect correlation to occur. In mathematical terms, Weatherill et al. (2014) describe spatial correlation as the covariance of the intra-event residuals at all sites, which are independent of the earthquake itself. Many studies have been conducted on spatial correlation (e.g. Wesson and Perkins, 2001; Boore et al., 2003; Wang and Takada, 2005; Lee and Kiremidjian, 2007; Goda and Hong, 2008; Goda and Atkinson, 2009; Jayaram and Baker, 2009; Foulser-Piggott and Stafford, 2010; Esposito and Iervolino, 2011), with the objective of producing a covariance matrix, which is populated by the spatial correlation co-efficients for all pairs of sites in the study area. The spatial correlation co-efficients are determined by empirically fitting ground shaking intensity records from past events to a statistical model. Some use single significant events such as the 1994 Northridge earthquake (Lee and Kiremidjian, 2007; Jayaram and Baker, 2009; Foulser-Piggott and Stafford, 2010) or the 2001 Chi-Chi earthquake (Jayaram and Baker, 2009; Foulser-Piggott and Stafford, 2010) as these are amongst the few events for which significant ground motion datasets exist. Other studies have used an ensemble of strong motion records from multiple events to produce more general spatial correlation models for Japan (Wang and Takada, 2005) and Europe (Esposito and Iervolino, 2011).

The correlation structure of a random variable can be represented by a semivariogram, which is a measure of the average dissimilarity between the values at two sites, u and $u+h$, based on their

separation distance, h (Goovaerts, 1997). The stationary semivariogram, γ, of the normalized intra-event residuals at each site, ε'_u and ε'_{u+h}, is therefore given by Eq. 2.3, (where $\varepsilon' = \varepsilon/\sigma$). The semivariogram measures the opposite of correlation, ρ, and the relationship between the two is given by Eq. 2.4, where a_{SC} is the sill, which is equal to the variance of the random variable, and which is 1 since the intra-event residuals have been normalised.

$$\gamma = \frac{1}{2}\left[E\left\{ \varepsilon'_u - \varepsilon'_{u+h} \right\}^2 \right] \tag{2.3}$$

$$\gamma = a_{SC}\left(1-\rho\right) \tag{2.4}$$

The observed stationary semivariogram for a specific value of h can be derived by calculating the average observed value of γ for all N_{SC} pairs of sites with separation distance h, as shown in Eq. 2.5. When repeated for all values of h, this gives an observed semivariogram plot to which a continuous semi-definite (non-negative) function is fit, so that the semivariogram and correlation can then be predicted for any value of h.

$$\widehat{\gamma} = \frac{1}{2N}\sum_{i=1}^{N_{SC}} \left\{ \varepsilon'_{u_i} - \varepsilon'_{u_i+h} \right\}^2 \tag{2.5}$$

When deriving the observed semivariogram, a balance has to be achieved between ensuring there are sufficient points with which to fit the model but also ensuring that each point is itself based on a sufficient number of observations to provide a reliable estimate of γ for that value of h. Esposito and Iervolino (2011) suggest that each point should be based on a minimum sample size of 30 and it may therefore be necessary to group the data into bins for h rather than using individual values of h. To fit the model, a functional form has to be assumed that fits the conditions of non-negativity and of $\gamma = 0$ for $h = 0$ and $\gamma = 1$ for $h = \infty$. For example, Boore et al. (2003), Wang and Takada (2005) and Jayaram and Baker (2009) all use variations of a single-parameter exponential model, while Goda and Atkinson (2009) adopt a three-parameter variant of the exponential model. Jayaram and Baker (2009) also test a Gaussian model and a spherical

model but only proceed with the exponential model for detailed analysis. These six functional forms are shown in Table 2.3. The least squares method can be used to determine the optimum values of the parameters for each functional form by minimising the sum of the squares of the errors. However, Jayaram and Baker (2009) state that the focus of the optimisation should be to ensure that the model fits well at short separation distances when the correlation is high. At greater distances, when the correlation is low, the effect that a site has on the residual of another site is likely to be negligible. Jayaram and Baker (2009) and Esposito and Iervolino (2011) therefore fit their models visually. Jayaram and Baker (2009) do not explicitly define a maximum value that can be considered to be a short separation distance but a visual inspection of their empirical semivariogram models for the 1994 M_W 6.7 Northridge earthquake in Los Angeles and the 1999 M_W 7.6 Chi-Chi earthquake in Taiwan, indicate that the models have been fit to optimise performance for site separation distances up to 15 km. Esposito and Iervolino (2011) optimise their model to fit distances up to 13.5 km.

Table 2.3 – Functional forms for spatial correlation models in the literature

Study	**Functional form**	**Eq. No.**
Boore et al. (2003):	$\gamma = 1 - \exp\left(-\alpha_B \sqrt{h}\right)$	(2.6)
Wang and Takada (2005):	$\gamma = 1 - \exp\left(-\alpha_{WT} h\right)$	(2.7)
Goda and Atkinson (2009):	$\gamma = 1 - \max\left\{\kappa_{GA} \exp\left[-\alpha_{GA} h^{\beta_{GA}}\right] - \kappa_{GA} + 1, 0\right\}$	(2.8)
Jayaram and Baker (2009): (exponential model)	$\gamma = 1 - \exp\left(\frac{-3h}{b_{JB}}\right)$	(2.9)
Jayaram and Baker (2009): (Gaussian model)	$\gamma = 1 - \exp\left(\frac{-3h^2}{b_{JB}^2}\right)$	(2.10)
Jayaram and Baker (2009): (spherical model)	$\gamma = \frac{3h}{2b_{JB}} - \frac{h^3}{2b_{JB}^3}$ (if $h \le b_{JB}$) $\gamma = 0$ (otherwise)	(2.11)

A method for simulating a field of spatially correlated intra-event residuals is described Weatherill et al. (2014) and is based on the Matrix Decomposition method proposed by Davis (1987). The final selected semivariogram model is used to calculate the correlation between each pair of sites according to their separation distance, *h*. The site to site correlations are then grouped into a matrix **C**, which is the covariance matrix. Since **C** is by definition positive-definite (it is symmetric and all its eigenvalues are positive), there exists a lower triangular matrix, **L**, such that $\mathbf{C} = \mathbf{L}\mathbf{L}^{T}$. The matrix **L** can be determined from **C** by Cholesky factorisation. The final vector of spatially correlated intra-event residuals, **Y**, is then calculated by Eq. 2.12, where $\boldsymbol{\mu}$ is the vector of mean intra-event residuals, which is a null vector, and **Z** is the vector uncorrelated intra-event residuals initially generated for each site.

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{L}\mathbf{Z} \tag{2.12}$$

An alternative method is to use the spatial correlation co-efficients to calculate joint ground shaking intensity exceedance probabilities at pairs of sites (Wang and Takada, 2005; Lee and Kiremidjian, 2007; Esposito and Iervolino, 2011). Whilst this approach is necessary for PSHA, it is not useful for the hybrid Monte Carlo simulation approach being used in this research, which does not consider ground shaking exceedance probabilities.

2.3.4 Ground deformation

Permanent ground deformations can be caused by three types of ground failure hazard: liquefaction, landsliding and surface fault rupture. The following sections describe existing methods for predicting permanent ground deformations caused by each hazard, with a focus on simplified methods that can be more easily applied in the insurance sector.

2.3.4.1 Liquefaction

Liquefaction is a phenomenon in which saturated soil loses strength due to excess pore water pressure generated by ground shaking. Not all soil is susceptible to liquefaction and it is more likely to occur in soils that are unconsolidated or uncompacted artificial fills. Sands and silty

sands are the soil types most susceptible to liquefaction but silt and gravel can also liquefy. Since these soil types generally exist close to the surface, high water tables also increase the risk of liquefaction. Liquefaction can manifest itself as either vertical settlement or horizontal lateral spread. Lateral spread occurs when liquefied soil is able to flow towards a free face such as a river.

The recent earthquakes in Haiti (2010), Canterbury, New Zealand (2010-11) and Tohoku, Japan (2011) highlighted the significance of liquefaction as a secondary hazard of seismic events and the significant damage that it can cause to buildings and infrastructure. However, the insurance sector was caught out by these events, with catastrophe models underestimating the extent and severity of liquefaction that occurred (Drayton and Verdon, 2013). A contributing factor to this is that the method used by some catastrophe models to account for liquefaction is based only on liquefaction susceptibility, a qualitative parameter that considers only surficial geology characteristics. Furthermore, losses arising from liquefaction are predicted by adding an amplifier to losses predicted due to building damage caused by ground shaking (Drayton and Vernon, 2013). There is a paucity of past event data on which to calibrate an amplifier and consequently, significant losses from liquefaction damage will only be predicted if significant losses are already predicted from ground shaking, whereas it is known that liquefaction can be triggered at relatively low ground shaking intensities (Quigley et al., 2013). Whilst liquefaction is less likely than ground shaking to be responsible for major building failures (Bird and Bommer, 2004), it can have a major impact on critical infrastructure such as roads, pipes and buried cables.

Bird and Bommer (2004) surmised that there are three options for dealing with ground failure hazards in loss estimation. They can be ignored; they can be modelled using a simplified approach; or they can be modelled by conducting a detailed geotechnical assessment. The first of these options will likely lead to underestimation of losses in earthquakes where liquefaction is a major hazard and lead to recurrence of the problems faced by insurers following the 2010-11 Canterbury earthquakes in particular. The last option, detailed assessment, is appropriate for

single-site risk analysis but is impractical for insurance loss estimation purposes because: 1) insurers are unlikely to have access to much of the detailed geotechnical data required as inputs to these methods; 2) they may not have the in-house expertise to correctly apply such methods and engaging consultants may not be a viable option; and 3) loss estimation studies are often conducted on a regional, national or supra-national scale for which detailed assessment would be too expensive and time-consuming. Therefore for the insurance sector, simplified approaches currently offer the most practical option for accounting for ground failures.

There are three stages to predict the occurrence of liquefaction and its scale (Bird et al., 2006). First it is necessary to determine whether soils are susceptible to liquefaction. Liquefaction susceptibility is based solely on ground conditions with no earthquake-specific information. This is often done qualitatively (Bird and Bommer, 2004) and currently this is also the full extent to which liquefaction risk is considered in some catastrophe models (Drayton and Verdon, 2013). The next step is to determine liquefaction triggering, which determines the likelihood of liquefaction for a given earthquake based on the susceptibility and other earthquake-specific parameters. Finally the scale of liquefaction can be predicted as a permanent ground deformation (PGDf).

Whilst the majority of methods for modelling liquefaction triggering required detailed site-specific geotechnical data, there are three methods that can potentially be used to predict liquefaction triggering and that only require data which are in the public domain or could be easily obtained without significant time or cost implications. These are the Liquefaction Potential Index (LPI) method (Iwasaki et al., 1984); the HAZUS method (FEMA, 2015); and the method developed by Zhu et al. (2015) that makes use only of remote sensing data that are publicly accessible from the United States Geological Survey (USGS) website. Furthermore, these methods are appropriate for regional-scale analysis and although some engineering judgment is required in their application, they do not require specialist geotechnical expertise, and as such they are appropriate for use by the insurance sector.

The LPI method (Iwasaki et al., 1984) is an extension of the factor of safety method developed by Seed and Idriss (1967). The basis of this framework is to compare two parameters: the cyclic stress ratio and the cyclic resistance ratio. The cyclic stress ratio depends on acceleration and shear stress and represents the load imposed on a site by an earthquake. The cyclic resistance ratio depends on geotechnical properties of the soil and represents the ability of the ground to withstand the load. The factor of safety against liquefaction is defined as the ratio of the cyclic resistance ratio and the cyclic stress ratio, for a layer of soil at depth, *z*. Liquefaction is predicted to occur if the factor of safety is less than 1. The cyclic stress ratio, CSR, is calculated by Eq. 2.13, where *g* is the acceleration of gravity; σ_v is the total overburden stress at depth *z*; σ'_v is the effective overburden stress at depth *z*; and r_d is a shear stress reduction coefficient given by Eq. 2.14 or 2.15.

$$CSR = 0.65\left(\frac{PGA}{g}\right)\left(\frac{\sigma_v}{\sigma_v'}\right)r_d \tag{2.13}$$

for z < 9.2 m: $$r_d = 1 - 0.00765z \tag{2.14}$$

for z ≥ 9.2 m: $$r_d = 1.174 - 0.0267z \tag{2.15}$$

Cyclic resistance ratio is normally calculated from geotechnical parameters based on cone penetration test (CPT) or standard penetration test (SPT) results. However, this data is rarely in the public domain and difficult to acquire on a regional scale. Alternatively, Andrus and Stokoe (2000) propose a method for calculating the cyclic resistance ratio, CRR, based on shear-wave velocity, V_S, as shown in Eq. 2.16, where V_{S1} is the stress-corrected shear wave velocity; V_{S1}^* is the limiting upper value of V_{S1} for cyclic liquefaction occurrence, which varies between 200-215 m/s depending on the fines content of the soil; and MSF is a magnitude scaling factor to account for earthquakes with moment magnitude other than 7.5. The advantage of this method is that whilst it also requires ground investigation, V_S can be measured by non-penetrative techniques, which may be easier in many locations. Furthermore, for the insurance industry the method may be advantageous if proxies for V_S can be obtained from publicly available datasets,

e.g. the United States Geological Survey (USGS) Global V_{S30} Map Server (USGS, 2013). V_{S1} is given by Eq. 2.17, where P_a is a reference stress of 100 kPa. The magnitude scaling factor is given by Eq. 2.18.

$$CRR = \left[0.022\left(\frac{V_{S1}}{100}\right)^2 + 2.8\left(\frac{1}{V_{S1}^* - V_{S1}} - \frac{1}{V_{S1}^*}\right)\right] \times MSF \tag{2.16}$$

$$V_{S1} = V_S\left(\frac{P_a}{\sigma_v'}\right)^{0.25} \tag{2.17}$$

$$MSF = \left(\frac{M_W}{7.5}\right)^{-2.56} \tag{2.18}$$

However Juang et al. (2005) found that Eq. 2.16 is conservative for calculating CRR, resulting in lower factors of safety and over-prediction of liquefaction occurrence. To correct for this, they propose a multiplication factor of 1.4, to obtain an unbiased estimate of the factor of safety, FS^*, given by Eq. 2.19. The value of the factor has been derived by comparing estimates of CRR using Eq. 2.16 to observations from 225 liquefaction case histories.

$$FS^* = 1.4 \times \frac{CRR}{CSR} \tag{2.19}$$

FS^* is an indicator of potential liquefaction at a specific depth. However, Iwasaki et al. (1984) noted that damage to structures due to liquefaction was affected by the severity of liquefaction at ground level and so propose an extension to the factor of safety method, the liquefaction potential index (LPI), which predicts the likelihood of liquefaction at surface-level by integrating a function of the factors of safety for each soil layer within the top 20 m of soil. The LPI at a site is calculated from Eq. 2.20, where $F^* = 1 - FS^*$ for a single soil layer. The soil profile can be sub-divided into any number of layers (e.g. twenty 1 m layers or forty 0.5 m layers), depending on the resolution of data available. Using site data from a collection of nine Japanese earthquakes between 1891 and 1978, Iwasaki et al. (1984) calibrated the LPI model and determined guideline criteria for determining liquefaction risk. These criteria propose that

liquefaction risk is very low for LPI = 0; low for 0 < LPI ≤ 5; high for 5 < LPI ≤ 15; and very high for LPI > 15.

$$LPI = \int_0^{20} F^* (10 - 0.5z)\, dz \tag{2.20}$$

Probability of liquefaction can also be used as a way of predicting liquefaction triggering, either by setting a threshold probability or by randomly sampling from a Bernoulli distribution with parameter equal to the probability of liquefaction. Hazus (FEMA, 2015) includes a simplified method for predicting the probability of liquefaction at a site, which only requires earthquake magnitude, PGA, water table depth and liquefaction susceptibility as inputs.

The liquefaction susceptibility is inferred from existing geological maps, with areas divided into six zones of liquefaction susceptibility, from 'None' to 'Very high', according to the classification of Youd and Perkins (1978). For a given liquefaction susceptibility category, the probability of liquefaction occurrence is given by Eq. 2.21, where $P[Liq \mid PGA=a_{max}]$ is the conditional probability of liquefaction occurrence for a given susceptibility zone at a specified level of peak horizontal ground motion, a_{max}, based on work by Liao et al. (1988); K_M is the moment magnitude correction factor, which accounts for the fact that the conditional probability relationships have been developed for a moment magnitude 7.5 earthquake; K_W is the ground water correction factor; and P_{ml} is the proportion of map unit susceptible to liquefaction, which accounts for the real variation in susceptibility across similar geologic units. The conditional probability and map unit factor are zero for the susceptibility zone 'None'. For the other zones, the conditional probability is given by linear functions of PGA and the map unit factor is a constant value and these are shown in Table 2.4. The moment magnitude and ground water correction factors are given by Eq.'s 2.22 and 2.23, where d_w is the depth to ground water. The map unit factor is a constant for each susceptibility zone, with values of 0.25, 0.20, 0.10, 0.05, 0.02 and 0, going from 'Very high' to 'None'.

$$P\left[Liq\right] = \frac{P\left[Liq \mid PGA = a_{\max}\right]}{K_M K_W} P_{ml} \tag{2.21}$$

$$K_M = 0.0027{M_W}^3 - 0.00267{M_W}^2 - 0.2055M_W + 2.9188 \tag{2.22}$$

$$K_W = 0.022d_w + 0.93 \tag{2.23}$$

Table 2.4 – Zone varying parameters for HAZUS method for predicting liquefaction probability (excluding 'None' zone)

Liquefaction susceptibility zone	**Conditional liquefaction probability, *P*[*Liq* | *PGA*=a_{max}]**	**Map unit factor, P_{ml}**
Very high	$9.09a_{max} - 0.82$	0.25
High	$7.67a_{max} - 0.92$	0.20
Medium	$6.67a_{\max} - 1.00$	0.10
Low	$5.57a_{max} - 1.18$	0.05
Very low	$4.16a_{max} - 1.08$	0.02

Zhu et al. (2015) also propose a simplified method for predicting the probability of liquefaction. The focus of their study is the development of prediction functions for use in rapid response and loss estimation, so they deliberately use predictor variables that can be calculated or are in the public domain and do not require any specialist knowledge to be applied. The input data include earthquake magnitude, PGA, V_{S30}, which is the average shear wave velocity across the top 30 m of soil, the compound topographic index (CTI), which is a proxy for saturation and the normalised distance (ND), which is the ratio between the distance to the coast and the distance between the coast and inland edge of the sedimentary basin (soil/rock boundary). Global estimates for V_{S30} at approximately 674 m grid intervals are available as open-access from the web-based USGS Global V_{S30} Map Server (USGS, 2013). CTI can also be obtained globally from the USGS Earth Explorer web service (USGS, 2014) and to calculate ND, the location of the inland edge can be estimated from a surface roughness calculation based on a digital elevation model (USGS, 2014), or by using V_{S30} data such that the inland edge is assumed to be

the boundary between NEHRP site classes C (soft rock) and D (stiff soil), (i.e. at V_{S30} = 360 m/s). Zhu et al. (2015) propose that the probability of liquefaction can be estimated as the logistic function given by Eq. 2.24, where X is a linear function of some set of the predictor variables.

$$P\left[Liq\right] = \frac{1}{1+e^{-X}} \tag{2.24}$$

Zhu et al. (2015) propose three linear models that are applicable to Christchurch and are tested in this study: a specific local model derived for Christchurch; a regional model for use in coastal sedimentary basins (including Christchurch) and a global model that is applicable more generally. For the global model, the linear predictor function, X_G, is given by Eq. 2.25, where the magnitude weighting factor, MWF, is given by Eq. 2.26. For the regional model, the linear predictor function, X_R, is given by Eq. 2.27 and for the Christchurch-specific local model, the linear predictor function, X_L, is given by Eq. 2.28.

$$X_G = 24.1 + \ln\left(PGA \times MWF\right) + 0.355CTI - 4.784\ln V_{S30} \tag{2.25}$$

$$MWF = \frac{M_W{}^{2.56}}{10^{2.24}} \tag{2.26}$$

$$X_R = 15.83 + 1.443\ln\left(PGA \times MWF\right) + 0.136CTI - 9.759ND - 2.764\ln V_{S30} \tag{2.27}$$

$$X_L = 0.316 + 1.225\ln\left(PGA \times MWF\right) + 0.145CTI - 9.708ND \tag{2.28}$$

Bird et al. (2006) list all the existing methods for predicting the scale of liquefaction in terms of permanent ground deformation (PGDf). The methods are divided into those that predict vertical settlement and those that predict lateral spreading. As with liquefaction triggering, almost all of these methods require results from CPT or SPT tests and detailed geotechnical data e.g. fines content, median particle size diameter. The likelihood that insurers possess or are able to acquire such data is low, which means that these approaches are not suitable for regional-scale rapid assessment. The lack of simplified models (i.e. simplified in the context of commercial

insurance application) is not surprising given the small number of models that exist for liquefaction triggering assessment and that by definition measuring the scale of liquefaction is more complex.

Hazus (FEMA, 2015) is the only source of a simplified method for predicting PGDf due to vertical settlement and is consistent with work presented by Tokimatsu and Seed (1987), although Bird et al. (2006) note that this method is only accurate to within a factor 2 or 3. The vertical permanent ground deformation, $PGDf_V$, is calculated by Eq.2.29 where $E[PGDf_V]$ is the expected vertical ground deformation, which is sampled from a uniform distribution whose bounds are defined by a characteristic amplitude, such that the lower bound is half the characteristic amplitude and the upper bound is double the characteristic amplitude. The characteristic amplitude varies by liquefaction susceptibility zone and these are shown in Table 2.5.

$$PGDf_V = P\left[Liq\right] \times E\left[PGDf_V\right] \tag{2.29}$$

Table 2.5 – Liquefaction susceptibility zone varying parameters for HAZUS method for predicting permanent ground deformations (excluding 'None' zone)

Liquefaction susceptibility zone	**Liquefaction triggering threshold, PGA(t) (g)**	**$PGDf_V$ characteristic amplitude (m)**
Very high	0.09	0.305
High	0.12	0.152
Medium	0.15	0.051
Low	0.21	0.025
Very low	0.26	0.000

Simplified methods for predicting PGDf due to lateral spread are proposed by Youd and Perkins (1987), Rauch and Martin (2000) and Hazus (FEMA, 2015), although the Hazus method is an extension of the Youd and Perkins method, so the latter is not discussed further. In Hazus, horizontal permanent ground deformation, $PGDf_H$, is calculated by Eq. 2.30, where K_Δ is a

displacement correction factor, given by Eq. 2.31, and the additional term on the right of Eq. 2.30 is the expected ground deformation for a given liquefaction susceptibility zone, which is a function of the normalized peak ground acceleration, $PGA_{norm} = PGA / PGA(t)$, where $PGA(t)$ is the liquefaction triggering threshold acceleration for that zone, as shown in Table 2.5.

$$PGDf_H = K_\Delta \times E\left[PGDf_H \mid \left(PGA / PL_{SC} \right) = a \right] \tag{2.30}$$

where,

$$K_\Delta = 0.0086M^3 - 0.0914M^2 + 0.4698M - 0.9835 \tag{2.31}$$

for $1 < PGA_{norm} < 2$: $$E\left[PGDf_H \mid \left(PGA / PL_{SC} \right) = a \right] = 12 \times PGA_{norm} - 12 \tag{2.32}$$

for $2 < PGA_{norm} < 3$: $$E\left[PGDf_H \mid \left(PGA / PL_{SC} \right) = a \right] = 18 \times PGA_{norm} - 24 \tag{2.33}$$

for $PGA_{norm} > 3$: $$E\left[PGDf_H \mid \left(PGA / PL_{SC} \right) = a \right] = 70 \times PGA_{norm} - 180 \tag{2.34}$$

The study by Rauch and Martin (2000), known as EPOLLS, propose a model where $PGDf_H$ is given by Eq. 2.35, where R_f is the shortest horizontal distance to the surface projection of the fault rupture, and T_d is the duration of ground motion between the first and last occurrence of accelerations ≥ 0.05 g at each site.

$$PGDf_H = \left(0.613M_W - 0.0139R_f - 2.42PGA - 0.01147T_d - 2.21 \right)^2 + 0.149 \tag{2.35}$$

2.3.4.2 Landslides

Landslides occur when the combination of static and inertia forces lead to instability in a mass and cause that mass to slide. With respect to earthquakes, this instability occurs when PGA exceeds a critical value, known as the critical acceleration. Two methods for predicting permanent ground PGDf associated with landslides are proposed by Hazus (FEMA, 2015) and Saygili and Rathje (2008). Both of these methods require the estimation of the critical acceleration value. However, currently there are no generally accepted or simplified methods for estimating the critical acceleration at a site. In the Hazus model, sites are classified into one of

ten landslide susceptibility zone based on the ground type (rock, soft rock or clay/silty soil), the groundwater level and the slope angle. A single value for critical acceleration is defined for each susceptibility zone. The permanent ground deformation, $PGDf_{LAND}$, is then given by Eq. 2.36, where PGA is the observed peak ground acceleration, n is the number of cycles, given by Eq. 2.37, and $E[d/PGA]$ is the displacement factor, which is a function of the ratio of the critical acceleration to PGA, although the functional form is not published.

$$PGDf_{LAND} = E\left[d / PGA\right] \times PGA \times n_{LAND} \tag{2.36}$$

$$n_{LAND} = 0.3419M_W^3 - 5.5214M_W^2 + 33.6154M_W - 70.7692 \tag{2.37}$$

In the empirically derived scalar model by Saygili and Rathje (2008), a median estimate of $PGDf_{LAND}$ is given by Eq. 2.38, where a_c is the critical acceleration; PGA is the observed peak ground acceleration and ε_{LAND} is a normally distributed random variable with zero mean and unit standard deviation.

$$\begin{aligned} \ln\left(PGDf_{LAND}\right) = {} & 5.52 - 4.43\left(\frac{a_c}{PGA}\right) - 20.39\left(\frac{a_c}{PGA}\right)^2 \\ & + 42.61\left(\frac{a_c}{PGA}\right)^3 - 28.74\left(\frac{a_c}{PGA}\right)^4 \\ & + 0.72\ln\left(PGA\right) + 1.13\varepsilon_{LAND} \end{aligned} \tag{2.38}$$

Since the equation for the displacement factor in the Hazus method is not published, only the Saygili and Rathje (2008) method is viable. However, this requires the estimation of critical acceleration, which can be achieved from the Hazus method, but only if information about geological conditions and slope angle are available.

2.3.4.3 Surface fault rupture

The only existing simplified method for predicting permanent ground deformations due to surface fault rupture, $PGDf_{RUP}$, is an empirical function proposed by Wells and Coppersmith (1994), which applies to all styles of faulting. This relates the median estimate of the maximum

permanent ground deformation along the fault, $PGDf_{RUPMAX}$, to the moment magnitude, M_W, as shown in Eq. 2.39.

$$\ln\left(\overline{PGDf_{RUPMAX}}\right) = -5.26 + 0.79M_W \tag{2.39}$$

The maximum deformation can occur at any point along the rupture and is assumed to drop linearly to zero at either end. To account for uncertainty in the model, $PGDf_{RUPMAX}$ is estimated from a uniform distribution between 0.5 and 2 times the median estimate.

2.4 Fragility

The purpose of the vulnerability module of an earthquake catastrophe model is to estimate the level of asset damage expected for different levels of earthquake intensity (Mahdyiar and Porter, 2005). In commercial catastrophe models this is usually achieved using vulnerability curves that relate damage ratio (the ratio of expected repair cost to replacement cost) to ground shaking intensity (LMA, 2013). More generally, a vulnerability curve is a function that relates a parameter of loss to an IM (Rossetto et al., 2014). However, for engineering purposes, it is more useful to relate the observed IM to a measure of damage, rather than directly to loss. A common tool used in the seismic risk assessment of buildings are fragility curves, which relate the probability of exceedance a defined damage state to a parameter of ground shaking (or ground deformation) intensity (Rossetto and Elnashai, 2003). In the case of critical infrastructure, it is necessary to know the expected physical condition of asset after an earthquake so that system performance can be measured and indirect losses can be estimated. Therefore, if indirect loss estimation is to be adequately addressed in catastrophe models, it must be through the application of fragility curves rather than vulnerability curves.

Pitilakis et al. (2014b) have documented all existing methods for estimating the fragility of assets in the systems considered by the SYNER-G study and it is observed that fragility curves are used for all nodal assets. Fragility curves can take a range of functional forms although Baker (2015) observes that the lognormal cumulative distribution is the most commonly

adopted. Other functional forms are possible however, such as binomial distribution (e.g. Ottonelli et al., 2015), beta distribution (e.g. Rossetto and Elnashai, 2003) or generalised linear model (e.g. Charvet et al., 2014). Rossetto et al. (2015) state that the popularity of the lognormal cumulative distribution is due to three reasons. Firstly, the y-axis is constrained between 0 and 1, which fits the probabilistic nature of fragility curves. Secondly, the x-axis is constrained between 0 and ∞, which fits the range of most IMs. Finally, the lognormal cumulative distribution is skewed to the left, which better reflects the clustering of data at low IM values. If a lognormal cumulative distribution is assumed, the fragility curve function takes the form shown in Eq. 2.40, where $P[d \geq d_i \mid \text{IM}]$ is the exceedance probability for the i^{th} damage state, IM is the ground motion intensity (e.g. PGA), Φ is the standard normal cumulative distribution function, β_{di} is the standard deviation of the natural logarithm (dispersion) of ground motion intensity for the i^{th} damage state and $\overline{IM_{d_i}}$ is the median value of ground motion intensity at which the asset is expected to reach the i^{th} damage state.

$$P\left[d \geq d_i \mid IM\right] = \Phi\left[\frac{1}{\beta_{d_i}} \ln\left(\frac{IM}{\overline{IM_{d_i}}}\right)\right] \tag{2.40}$$

If different damage states are defined for an asset, then distinct fragility curves exist for each damage state with different values for the dispersion, β_{di}, and for the median ground motion, $\overline{IM_{d_i}}$. Therefore the risk to an asset being assessed for multiple damage states will be represented by a set of fragility curves, as shown in Figure 2.2. Whilst the vertical axis of a fragility curve specifies the exceedance probability of a damage state, the probability of being in a specific damage state can be determined by taking the difference between the exceedance probabilities for that damage state and the damage state above, as shown in Eq. 2.41 and on Figure 2.2. For assets that can be classified by different typologies, each typology will have its own set of fragility curves.

$$P\left[d = d_i \mid IM\right] = P\left[d \geq d_i \mid IM\right] - P\left[d \geq d_{i+1} \mid IM\right] \tag{2.41}$$

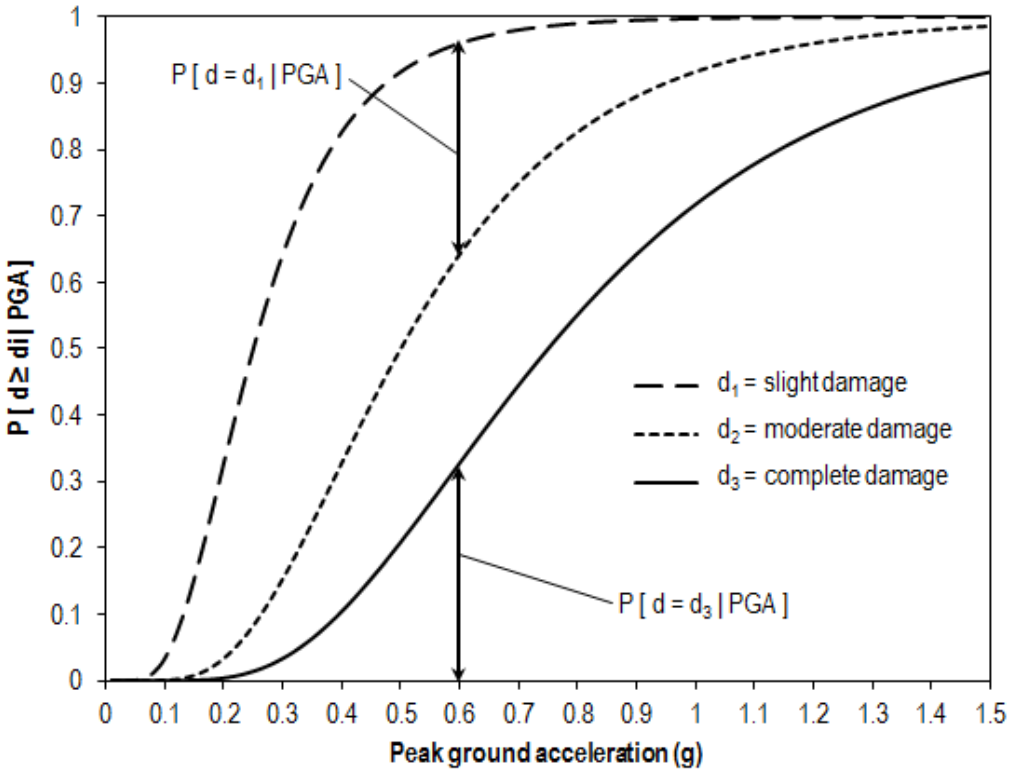

Figure 2.2 – Example set of fragility curves for an asset with three damage states

2.4.1 Existing fragility curves

When conducting a seismic risk assessment there is a choice between either using an existing fragility curve for an asset or developing a new one. The advantage of using existing curves is that the process is faster and does not require any specialist expertise to implement and therefore this is the most practical option for insurers. When selecting an existing fragility curve, the intention is to choose one for an asset that, in physical terms, matches as closely as possible the assets being assessed. Other factors to consider are the IM used by the fragility curve and whether it is one that is being predicted in the assessment and also the physical attributes used to distinguish between different asset typologies. In some cases the attributes necessary to apply a fragility curve from a particular study may not be known and so either the relevant attribute has to be assumed or an alternative curve has to be used. Pitilakis et al. (2014b) compiled a comprehensive database of all existing fragility curves for elements within the critical

infrastructure systems being studies in this research, with the exception of telecommunications. Based on their database, and supplemented by other sources found independently, Appendix C lists the references of known critical infrastructure asset fragility curves and the physical attribute data required to implement each one. A brief summary of this is provided in the following paragraphs.

In the electric power network, the focus of fragility studies has been on substations. There have been different approaches to this, with five studies producing fragility curves for substations as single entities (e.g. Vanzi, 2000; FEMA, 2015), and a further nine producing fragility curves for each internal components of a substation such as transformers, circuit breakers or switches. Whilst some of the component studies focus on key components only (e.g. Anagnos and Ostrom, 2000; Straub and der Kiureghian, 2008), others consider all components, which allows a fault tree analysis to be conducted to predict overall substation failure (e.g. Hwang and Huo, 1998; Rasulo et al., 2004). Voltage and seismic design standards are the most common classifying features. Hazus (FEMA, 2015) is the only study to produce fragility curves for generation plants with different sets of curves based on capacity and seismic design standards. Hazus (FEMA, 2015) also produces fragility curves for distribution circuits, although this is not the same as a conduit. A distribution circuit is defined as a grouping of all the conduits, supporting equipment and in-line equipment that feed off a single distribution substation. The damage state of the entire circuit can be predicted but it is not clear how the vulnerability of individual conduits can be predicted, which would be necessary for system performance analysis. Park et al. (2006) produce fragility curves for conduits based on data from the February 2001 M_W 6.8 Nisqually earthquake, but make no distinction between whether the conduits are overhead or buried, nor do they consider any other attribute.

For water supply and waste water system nodes, Hazus (FEMA, 2015) is the primary source of fragility curves, with curves published for sources (wells), treatment plants (water supply and waste water) and pumping/lift stations. Different curves are produced to account for different seismic design standards and in the case of treatment plants, for different capacities. Hazus

(FEMA, 2015) also publishes fragility curves for storage tanks although does not consider any typologies. More refined tank fragility curves are produced by O'Rourke and So (2000) and ALA (2001), which consider separate curves for attributes such as material and quantity of stored content. The SRM-LIFE study produces fragility curves for treatment plants and pumping stations, but this report is not in the public domain and its details are only available from the SYNER-G database (Pitilakis et al., 2014b). There are 25 studies that consider the fragility of pipes. A notable feature of all of these is that they do not in fact produce probabilistic fragility curves. All the methods are deterministic calculations of mean repair rate (RR), i.e. the mean number of repairs per unit distance. Repair rate functions can adopt a range of functional forms, although almost all those listed in Appendix C use a power relationship as shown in Eq. 2.42, where a_{RR} and b_{RR} are coefficients determined by regression.

$$RR = a_{RR} \times IM^{b_{RR}} \tag{2.42}$$

It is generally assumed that the distribution of intervals between repairs follows a Poisson distribution (Hwang et al. 1998; Adachi and Ellingwood 2008) and in such a case, the repair rate becomes the parameter of a Poisson process used to estimate the probability of a specified number of repairs occurring on a segment of pipe with length, L_{pipe} km (it is most common for repair rates to be expressed in terms of repairs per kilometre). Assuming that a pipe fails when it suffers at least one break, the Poisson process can be used to estimate the failure probability of the pipe as shown in Eq. 2.43 and Eq. 2.44.

$$P[\,fail\,] = P[repairs > 0] = 1 - P[repairs = 0] \tag{2.43}$$

$$P[\,fail\,] = 1 - e^{-RR \bullet L_{pipe}} \tag{2.44}$$

Most of the existing studies produce models to predict repair rates due to ground shaking with either PGA or PGV as IMs (although it is notable that the more recent studies tend to use PGV). Pineda-Parros and Najafi (2010) state that PGV should be most relevant to buried infrastructure since it relates more closely than PGA to ground strain. Furthermore, the studies by Isoyama

(2000) and O'Rourke et al. (2001) show very strong correlation between PGV and pipe damage. Seven of the studies produce models to predict repair rates due to ground deformation based on PGDf. Material is the most common attribute for distinguishing between pipe typologies, although to a lesser extent, diameter, soil type and joint type are also considered. There are six studies that propose fragility curves for tunnels based on construction type and ground conditions (e.g. ALA, 2001, FEMA, 2015), but only ALA (2001) propose fragility curves for canals.

For natural gas and oil delivery networks, fragility curves for pipelines are the same as for the water supply and waste water system pipes. The same applies to storage tanks, although there are two additional studies that produce fragility curves specifically for liquefied gas/oil tanks (Iervolino et al., 2004; Berahman and Behnamfar, 2007). For compressor stations, sources of fragility curves include LESSLOSS, SRM-LIFE, and Hazus (FEMA, 2015), although only Hazus, which identifies typologies by capacity and seismic design status, is in the public domain (information for LESSLOSS and SRM-LIFE is reported indirectly by Pitilakis et al. (2014b)).

For highways networks, only Hazus presents fragility curves for roads. These are with respect to PGDf and roads area classified by their importance. There are fourteen fragility functions for bridges and whilst in the past these were treated as single elements, more recently it has become common practice to treat bridges as systems of components each with their own fragility (Padgett and DesRoches, 2008). The most comprehensive study is Hazus (FEMA, 2015) which has produced fragility curves for 28 typical bridge classifications, based on seismic design status, number of spans, material, pier type, abutment type, bearing type and span continuity, while Moschonas et al. (2009), produced fragility curves for eleven typical bridges found in Greece, classified by the type of pier, the type of deck and the type of pier-deck connection. Zhang et al. (2008) have produced fragility curves for six bridge classifications in California based on span continuity and abutment type. Corrosion is also considered in some studies (Choe et al., 2009; Gardoni and Rosowsky, 2009) by applying a factor to the failure probability that

accounts for environmental, material and geometric properties that may make some bridges more susceptible to corrosion. Notably nine of the fourteen bridge fragility curves are only valid for a single specific typology. Fragility curves for tunnels are the same as for the water systems. Fragility curves also exist for embankments, slopes and retaining walls, although in practice an insurer would most probably focus on more easily identifiable assets such as roads, bridges and tunnels.

For telecommunications networks, only two studies present fragility curves. Hazus (FEMA, 2015) presents for curves central offices, whilst Leelardcharoen (2011) expands on this by producing separate curves for different types of central office. No fragility analysis has been undertaken for linear assets in telecommunications systems as it is generally thought that this type of infrastructure performs well in earthquakes (Schiff, 1997).

2.4.2 Development of new fragility curves

In some cases a fragility curve may not exist for the specific asset class being modelled, or even if one does exist, differences in construction practices may mean that curves developed for an asset class in one location may not be valid for the same asset class in another location. Using existing fragility curves for a 'similar' asset class may be convenient but can also be inaccurate. Therefore, if circumstances allow, the development of new curves is advantageous since it ensures that fragility analysis is based on a function that is specific to the assets being studied and their unique characteristics.

There are four ways new fragility curves can be developed: either empirically using data from past events; analytically using numerical methods and engineering analysis; by expert judgement; or by a hybrid of these methods, e.g. analytical methods used to enhance empirical data (Rossetto and Elnashai, 2003). Each of the methods possesses its own advantages and disadvantages as discussed by Rossetto and Elnashai (2003). Empirical functions are more straightforward to produce since they require only statistical analysis rather than specialist engineering knowledge but can be hindered by data deficiencies such as misclassification of

observed damage states, clustering of data in a narrow ground motion range or attributing damage to a single event that was actually caused by multiple earthquakes or separate phenomena (Rossetto et al., 2014). Analytical functions are potentially more reliable than empirical functions but require more computational effort, greater engineering expertise and challenges remain in terms of accounting for all building features. Expert judgment can also overcome some of the problems identified with empirical analysis but may suffer from the experience and subjectivity of the experts consulted.

State-of-the-art guidelines for the Global Earthquake Model (GEM) (Rossetto et al., 2014) list seven steps in the framework for empirical fragility curve development: preparation of damage data; selection and estimation of IM levels; selection of statistical models; fitting of statistical models; identification of optimum statistical model; validation of fragility curve; and reporting of the fragility curve. GEM guidelines for analytical fragility curve development (D'Ayala et al., 2013) list six steps: definition of index buildings; definition of components for response analysis; selection of model type (1D/2D/3D); definition of damage states; selection of model type (static or dynamic) and calculation of engineering demand parameter damage state thresholds; and constriction of fragility curves. However this should be conducted by an engineer with Master's level training and the ability to perform non-linear structural analysis.

2.4.3 Bayesian updating of existing fragility curves

In some cases there may be insufficient data with which to develop new completely new fragility curves for an asset but analysis shows that existing fragility curves are inadequate. In these cases an alternative solution is to adopt a Bayesian approach to the development of new fragility curves (e.g. Jaiswal et al., 2011; Rossetto et al., 2014). In the method proposed by Rossetto et al. (2014) for the Global Earthquake Model (GEM), the parameters of an existing fragility curve are used as prior estimates for the parameters of a new fragility curve developed from empirical data. The existing function parameters are considered to be random variables whose distribution is determined from observed data and some other prior knowledge. A prior

distribution is assumed for each parameter. Using Bayes' Theorem, the posterior distribution of the parameters (i.e. the distribution of the parameters for the new curve) are proportional to the assumed prior distribution (i.e. the distribution of parameters for the existing curve) and the likelihood function, as shown in Eq. 2.45, where for a set of parameter values, $\boldsymbol{\theta}_{FRAG}$, the likelihood function is the probability that the set of empirical observed damage states, $\boldsymbol{\alpha}_{FRAG}$, and corresponding IM observations, $\boldsymbol{\beta}_{FRAG}$, would occur given $\boldsymbol{\theta}_{FRAG}$.

$$f\left(\boldsymbol{\theta}_{FRAG} \mid \boldsymbol{\alpha}_{FRAG}, \boldsymbol{\beta}_{FRAG}\right) \propto f\left(\boldsymbol{\theta}_{FRAG}\right) L\left(\boldsymbol{\theta}_{FRAG}; \boldsymbol{\alpha}_{FRAG}, \boldsymbol{\beta}_{FRAG}\right) \tag{2.45}$$

Rossetto et al. (2014) propose using a Gibbs sampling method (Kruschke, 2011) to estimate the posterior distribution. The Gibbs sampling method is a type of Markov Chain Monte Carlo algorithm applicable to multivariate probability distributions. Initial values of each parameter are assumed from the prior distributions and the likelihood function is calculated to give a first estimate of the posterior distribution for each parameter. The Gibbs sampling algorithm subsequently generates a chain of sampled parameter values, where each sample realisation is dependent on the previous one. The likelihood function and posterior distribution estimates are calculated at each sample and after a large number of samples there should be convergence in the results. To allow for this, Rossetto et al. (2014) suggest that chains of 100,000 samples are used. The Bayesian method is particularly useful when an observed dataset is small or when the specific objective is to improve a function with newly acquired data.

2.4.4 Damage assignment

Fragility curves estimate the probability of an asset being in a specified damage state or worse, while repair rate functions can be manipulated to estimate the failure probability of a pipe. There are two common methods in the existing literature for deterministically predicting damage in a particular scenario with observed IMs. The first method evaluates the number or proportion of assets in each damage state from amongst a group of assets of the same typology and hence modelled with the same fragility curve (e.g. Kircher et al., 1997; King et al., 2004; Tang and Wen, 2009). The assets are grouped into bins according to their corresponding

observed IM. All the assets within in a bin are assumed to have the same IM value (usually the midpoint of the bin) and exceedance probabilities for each damage state are calculated for that assigned IM. For the assets in that bin, the proportion that are predicted to be in each damage state can be calculated from Eq. 2.41. This is repeated for all bins to evaluate the total count.

The second method is to assign a damage state to each asset individually. For each asset the damage state exceedance probabilities are evaluated from the fragility curve based on the observed IM at its location. The asset is then assigned into a specific damage state by sampling from a uniformly distributed random variable between 0 and 1. The process for assigning the damage state is that if the random variate is smaller than the exceedance probability for the worst damage state (i.e. complete damage or failure), then the asset is assigned to that damage state. If not, the random variate is compared to the exceedance probability for the next worst damage state and if it is smaller, the asset is assigned to that damage state. The process is repeated for all damage states, going in order of decreasing severity. For critical infrastructure system performance measurement, it is important to know not just how many assets are in a damage state, but specifically which ones, since the systemic impact of a failure will vary depending on the criticality of the failed asset (e.g. how many customers it serves, levels of redundancy). Therefore this method is much more prevalent in existing seismic risk assessments of critical infrastructure systems (e.g. Dueñas-Osorio et al., 2007; Wang et al., 2010; Poljansek et al., 2012).

2.5 System performance

The performance of infrastructure, also known as residual functionality (Cimellaro, 2013), can be measured in three different ways (Chang and Song, 2007). The first is to assess the performance of individual assets in the system as described in the previous section; the second is to assess whether there is a connection between two points in the system, (i.e. connectivity, e.g. Dueñas-Osorio et al., 2007; Kang et al., 2008; Bensi et al., 2009); and the third is to assess the quality of service, (i.e. serviceability, e.g. Chang and Nojima, 2001; Adachi and Ellingwood,

2009; Romero et al., 2010). Whilst asset performance is informative for estimating direct economic losses, the latter two perspectives are more useful for estimating indirect losses, since they tell us more about how the level of service has been affected to customers.

Connectivity analysis determines whether there is a functioning system connection between a source node and a demand node and only requires knowledge of whether the components that exist on the paths between the two nodes are functional. However, fragility curves for assets only predict a failure probability, not a functional state, and so given a set of IMs for a single event, there is uncertainty relating to the actual functional state of each asset. For a particular group of assets, fragility curves can estimate how many amongst that group will fail or be in a specific damage state. However, it is not just how many assets are in a particular damage state that is important but specifically which ones. There are two reasons for this. Firstly, systems can exhibit redundancy, which means that overall performance does not necessarily relate to quantity of failed assets. The geographical distribution of damaged components needs to be known to determine whether redundancy will mitigate the effects of damage. Secondly, the size and importance of areas served by assets can vary, so the system performance can vary considerably depending on which particular assets have failed.

Many connectivity analysis studies have analysed lifeline systems in a binary fashion (e.g. Chang and Song, 2007; Kang et al., 2008; Bensi et al., 2009). Every node or link is either in a 'functional' state or in a 'failed' state. If it is in a 'functional' state, then it is implicitly assumed to be functioning at 100% of its capacity. In reality however, some components may only be partially functioning. For example, a water pipe that has a number of leaks will still be able to supply water, but at reduced levels. Such a performance reduction is not measured in connectivity analysis, which is only concerned with whether or not a component is able to provide a service, not the level of service.

In the binary connectivity analysis, a single earthquake is modelled and the functional state of each asset has to be assigned. This is done by a random sampling procedure based on the failure

probability of the asset (i.e. fragility) and the result is a randomly generated realisation of damage to all the system assets. The performance of the whole system (connectivity or serviceability) is then computed for that system damage realisation. To capture the uncertainty in damage assignment, Monte Carlo simulation can be applied. If only a single scenario event is being studied, then the process is repeated by re-sampling the damage states to create different system damage realisation and by re-calculating the performance for each realisation. If multiple events are being studied, then as long as the number of events is large, one can expect that the uncertainty in damage assignment for a particular asset will be reflected across all event simulations. Consequently, it is not necessary to consider more than one system damage realisation per event. However, if only a small number of events are being studied, then it may be necessary to create multiple system damage realisations per event to ensure the uncertainty is adequately reflected.

An alternative to the binary approach to connectivity analysis is to estimate the probability that a connection exists between two assets. This is often referred to as an analytical approach as it does not require repeated sampling, and in theory allows for more rapid computation (Kim et al., 2009). A common analytical method is the Matrix System Reliability method (e.g. Chang and Song, 2007; Kang et al., 2008; Kim et al., 2009; Song and Kang, 2009; Song and Ok, 2010; Kang et al., 2012). To apply this method it is necessary to identify all possible system damage states and all the critical minimum paths between relevant node pairs (e.g. between sources node and demand nodes). The probability of disconnection between a node pair can be calculated from the failure probabilities of the assets along the minimum paths. For a single earthquake, there is no need to generate multiple system damage realisations to account for the uncertainty arising from the fragility curves, since failures are considered probabilistically. However, for loss estimation it is still necessary to model multiple earthquake scenarios to account for uncertainty in the hazard model. Therefore, whilst analytical methods are useful for single scenario event assessment, their primary benefit becomes redundant in multi-event risk assessment. Another disadvantage is that analytical methods can have problems dealing with

large-scale systems since the computational effort increases exponentially as the number assets in the system increases (Song and Ok, 2010). The problem of identifying all network states and minimum paths can become intractable to the extent that it becomes necessary to use heuristic sampling techniques, which negates the advantage of using analytical methods over simulation methods.

System performance metrics used in connectivity analysis can be a single global index for the entire system (e.g. Adachi and Ellingwood, 2008; Romero et al., 2010; Poljansek et al., 2012) or a local index for each individual demand asset. The major advantage of a global index is simplicity, as it allows a quick comparison to be made between different scenarios. However, a global index does not allow for the identification of which geographical areas within a system have problems. The value of global or local indices will depend on the needs of the assessment. If losses are being estimated for a small number of assets (or clusters of assets), then the geographical precision of local indices may be more beneficial. However, if the purpose of the assessment is to predict more general region-wide impacts, then the simplicity a global index might be more useful. An example of a local index for connectivity analysis is the connectivity loss which can be estimated for each demand node. (e.g. Duenas-Osorio et al., 2007; Adachi and Ellingwood, 2008; Kim et al., 2009; Poljansek et al., 2012; Cavalieri et al., 2014a; Esposito et al., 2015). The connectivity loss is the ratio of the number of source nodes to which the demand node is connected after an earthquake to the number of source nodes to which it is connected normally. Poljansek et al. (2012) also propose taking the mean of the connectivity loss at each demand node as a global index. One problem with connectivity loss as a metric is that although it has a mathematical definition, its practical meaning is unclear to a non-expert. Furthermore, it is possible for a system to have a high connectivity loss yet all customers are still connected to the system. Conversely, a system can have some customers disconnected yet its connectivity loss can be low. Since no population weighting is applied in the calculation of connectivity loss, it does not account for the fact that the social impact of loss of connectivity will vary between demand nodes. A more straightforward and meaningful global index for hierarchical systems is

to simply count the number of customers connected to a system after an earthquake (e.g. Reed et al., 2009) and this is often how system performances are reported after real earthquakes (e.g. Kameda, 2000; Ministry of Health, 2012; Orion, 2014).

Serviceability analysis determines whether a customer has a functioning system connection and also the quality of service the customer can expect to receive (e.g. is water at adequate pressure). Each infrastructure system has its own unique metrics, and therefore methodology, for measuring quality of service that depend on its components and the flow characteristics of the utility. However, the need to understand flow makes this more difficult for an insurer apply than connectivity analysis. As serviceability analysis allows consideration of partial functionality of assets, it can provide a better assessment of losses. Overall, very few studies exist that adopt serviceability analysis, and those that do exist tend to simply attribute a functionality level based on a damage state inferred from the fragility curve (e.g. ALA, 2001; Leelardcharoen, 2011).

For electric power networks standard power flow equations can be used to calculate the power supplied to each node and the demand at each node. Ang et al. (1996) suggest a range of local serviceability metrics that can be calculated at each node based on this data and a global metric which compares demand to supply for the system as a whole. Hydraulic analysis can be used to determine volume, pressure and head of water passing through each node of water supply, waste water and oil delivery systems (Rasulo et al., 2008), and equivalent pressure and flow analysis for gas networks can be conducted using the method described by Osiadacz (1987). For these systems a common metric is the system serviceability index (e.g. Vanzi, 1996; Adachi and Ellingwood, 2008; Javanbarg and Takada, 2009; Romero et al., 2010; Wang et al., 2010; Poljansek et al., 2012, Cavalieri et al., 2014a; Esposito et al., 2015), which is equivalent to the connectivity loss but instead of counting the number of sources connected to a demand node, it measures the quantity of the utility being supplied to each demand node in terms of volume or flow. For telecommunications networks Leelardcharoen (2011) proposes the blocking probability as a system performance metric. It is the probability that a call cannot be connected

because of system capacity constraints and can be calculated for a single node or for the network as a whole. Examples of system performance metrics for highways networks include length of highways open (Chang and Nojima, 2001), shortest path distance between node pairs (Chang and Nojima, 2001), demand to capacity ratio (Kawakami, 2000), average travel time delay (Kiremidjian et al., 2007; Jayaram and Baker, 2010)

System performance can be analysed as either steady-state or transient. Steady-state analysis takes a snapshot of system operation in equilibrium at a discrete moment in time, whereas transient analysis shows how the operation varies continuously over a time period. Ang et al. (1996) and Vanzi (1996) both observe that for electric power systems, the computational effort required for transient serviceability analysis is disproportionately large and so recommend the use of steady-state analysis. Whilst there is value in assessing the change in system performance over time from the earthquake to full restoration, transient analysis increases complexity and is computationally very intensive, and so alternative methods may be more appropriate for loss estimation. After an earthquake, improvements to system performance are likely to be step-like rather than continuous, since the completion of individual restoration activities would occur at discrete time intervals. Improvement of system performance can therefore be modelled as a series of steady-state analyses at discrete time-steps rather than continuously and there are methods in the literature to estimate the restoration time of assets (e.g. FEMA, 2015; Cagnan et al., 2006) that can be used to determine the time-steps. As well as measuring the performance at each time-step, the performance of a system can also be measured as a single value that accounts for outage duration. In the utilities sector, temporal system performance is often measured in people minutes, which is the total sum of the outage times for each individual customer (Orion, 2009).

There are only a very limited number of studies that have previously attempted to validate infrastructure system performance models with observed metrics from past earthquakes. Wang and O'Rourke (2006) and Bonneau and O'Rourke (2009) validate models of the Los Angeles water supply system with observations from 1994 Northridge earthquake. Wang and O'Rourke

(2006) measure performance in terms of flow rates at key locations and find simulation errors of up to 10%. Bonneau and O'Rourke (2009) measure the proportion of customers who remain connected to the system and predict a connection level of 71.4%, which is in the range of 70-75% estimated by the system operator. Javanbarg and Takada (2009) validate a model of the Osaka water supply system with measurements of water pressure at 20 nodes after the 1995 Kobe earthquake. Errors are not reported and results are only presented as plots, but it can be seen that in some cases errors exceed 50%. Wu and Dueñas-Osorio (2013) validate models of the electric power network and water supply system in Concepcion, Chile with observations from the 2010 Maule earthquake. While other studies aim for a mean prediction that is as close as possible to the observation, Wu and Dueñas-Osorio (2013) set a different objective of ensuring that the model does not underestimate performance. The performance metric is connectivity loss, which measures the average percentage reduction in the number of topological paths between each generation node and demand node. The calibrated model has a probability of 0.96 for predicting an electric power network performance at least as bad as observed and a corresponding probability of 0.72 for the water supply system.

Two other system performance validation studies exist in relation to hurricanes. Winkler et al. (2010) validate a model of the Harris County, Houston electric power network with observation from Hurricane Ike in 2008. The performance metric is the proportion of customers experiencing power outages and is measured separately for each of the 137 zip codes in the county. The mean absolute error is reported as 15.59%. Nateghi et al. (2011) validate five candidate models for an unspecified part of the Gulf Coast electric power network with observations from Hurricane Ivan in 2004 and from Hurricane Katrina and Hurricane Dennis in 2005. The performance metric is the mean outage duration time. For Hurricane Ivan – the event that is used for model development – the best performing model has an error of 16.4%. For Hurricane Katrina the relative error is 19.9% and for Hurricane Dennis the error is 27%.

2.6 Interdependencies

2.6.1 Characteristics of interdependency

It is possible to make a semantic distinction between dependency and interdependency when analysing infrastructure systems (e.g. Rinaldi et al., 2001; Bigger et al., 2009). Dependency describes a unidirectional relationship in which one system depends on another, whereas interdependency describes a bidirectional relationship where each system depends on the other. However, dependency can also be thought of as simply a specific case of interdependency in which the relationship is null in one direction and it is common in natural hazards literature to use the term 'interdependency analysis' to describe the analysis of any infrastructure system interaction irrespective of directionality (e.g. Chang et al., 2007; Dudenhoeffer et al., 2007, Adachi and Ellingwood, 2008; Reed et al., 2009; Poljansek et al., 2012). The general use of the term 'interdependency' is adopted here.

Rinaldi et al. (2001) identify six dimensions for describing an interdependent infrastructure relationship: infrastructure characteristics, environment, state of operation, type of interdependency, type of failure and response behaviour. Infrastructure characteristics describe the principal spatial, temporal, operational and organisational factors that affect how an infrastructure system can be expected to respond to a perturbation. Environment refers to the economic, political, legal, social and security context within which infrastructure systems operate. State of operation refers to current demand levels for an infrastructure system, which may impact on the ability of the system to absorb loss of capacity.

Rinaldi et al. (2001) summarised four main types of interdependent relationship: physical, geographical, cyber and logical. Physical interdependency refers to operational links between two systems in which the state of one system is dependent on the output of another, e.g. water supply system requires electric power to operate pumps. Geographical interdependency refers to the co-location of elements across two systems such that damage to one is likely to cause physical damage to the other, e.g. a pipe traversing a bridge could potentially withstand

earthquake forces but suffer damage due to collapse of the bridge. Cyber interdependency refers to systems connected by informational links; and logical interdependency refers to all other relationships not categorised by the other three types. Dudenhoeffer et al. (2007) propose a fifth type – social interdependency, defined as behaviour linked by more abstract factors such as public opinion, confidence, fear or culture. These are different to other 'logical' interdependencies which are policy-based or procedural. Duenas-Osorio and Kwasinski (2012) propose the concept of logistical interdependency where there is no physical link between systems but where damage to one system can hinder the repair and restoration of another system. An example of a logistical interdependency is the impact that damage to highways and telecommunications can have on the ability of other infrastructures to implement repair and restoration works.

An infrastructure element can fail after an earthquake in two ways. One is that the element may suffer physical damage as a direct result of seismic forces. If elements in different systems fail in this way due to the same earthquake, this is an example of common cause failure (Rinaldi et al., 2001) or inherent failure (Han and DeLaurentis, 2013). However an element may be able to withstand the initial shock yet stop functioning because it is no longer receiving some input from another system which it requires to operate normally, e.g. when a water pump loses its power supply. Han and DeLaurentis (2013) refer to this as a propagating failure, although Rinaldi et al. (2001) make a distinction between the case where failure in one system causes failure in a second system – cascading failure – and the case where failure in one system exacerbates an existing failure in a second system – escalating failure. In terms of seismic risk to infrastructure, fragility analysis, as described in section 2.4, predicts the occurrence of inherent failures in individual infrastructure assets. To correctly model system performance however, it is also necessary to account for propagating failures resulting from the different types of interdependency.

The final dimension of interdependency is response behaviour, which describes the coupling between systems and how the performance of one system can affect that of another. The study

of interdependencies in this thesis focuses specifically on this dimension for two purposes: the simulation of interdependencies to predict system performance after a shock (also referred to as interdependency simulation) and the quantification of the degree of interdependency between systems for strategic risk management.

2.6.2 Interdependency simulation

There are many different techniques that can be utilised to simulate lifelines interdependencies (Pitilakis and Kakderi, 2011) but collectively they fall into two paradigms, described by Oliva et al. (2010) as decomposed and monolithic approaches and by Eusgeld et al. (2011) as the integrated and coupled approaches. The primary distinctions between these two paradigms relate to the scale at which interdependency is modelled and the stage at which interdependencies are considered in the analysis. The differences in approach are summarised schematically in Figure 2.3 for a fabricated pair of lifeline systems: an electric power network and a water supply network. In this example, there is interdependency between the systems because a pumping station in the water supply network relies on electric power.

In the integrated (or decomposed) approach the interdependency is defined at the asset level. This means the specific assets between which the interdependency exists are identified and represented in a single multi-system network model. In the schematic in Figure 2.3, this is represented by the dashed line between the pumping station and the substation on which it is reliant for power. Defining these specific interfaces integrates the individual lifelines together into one single system, within which each lifeline becomes a sub-system. Once the interdependencies are identified, inherent failures are identified using fragility functions (an interdependency link may also have a fragility associated with it if it is representative of physical equipment) and then propagated through the integrated system. In the trivial example in Figure 2.3, the propagation of failures is straightforward, i.e. all elements downstream of the inherent failures are assumed to be non-operational unless they have a back-up or secondary input. However in more complex exercises, such as where more than two lifelines are being

assessed or where there is bidirectional interdependency between two lifelines, there is a possibility of an inherent failure propagating through a feedback mechanism back into its own system.

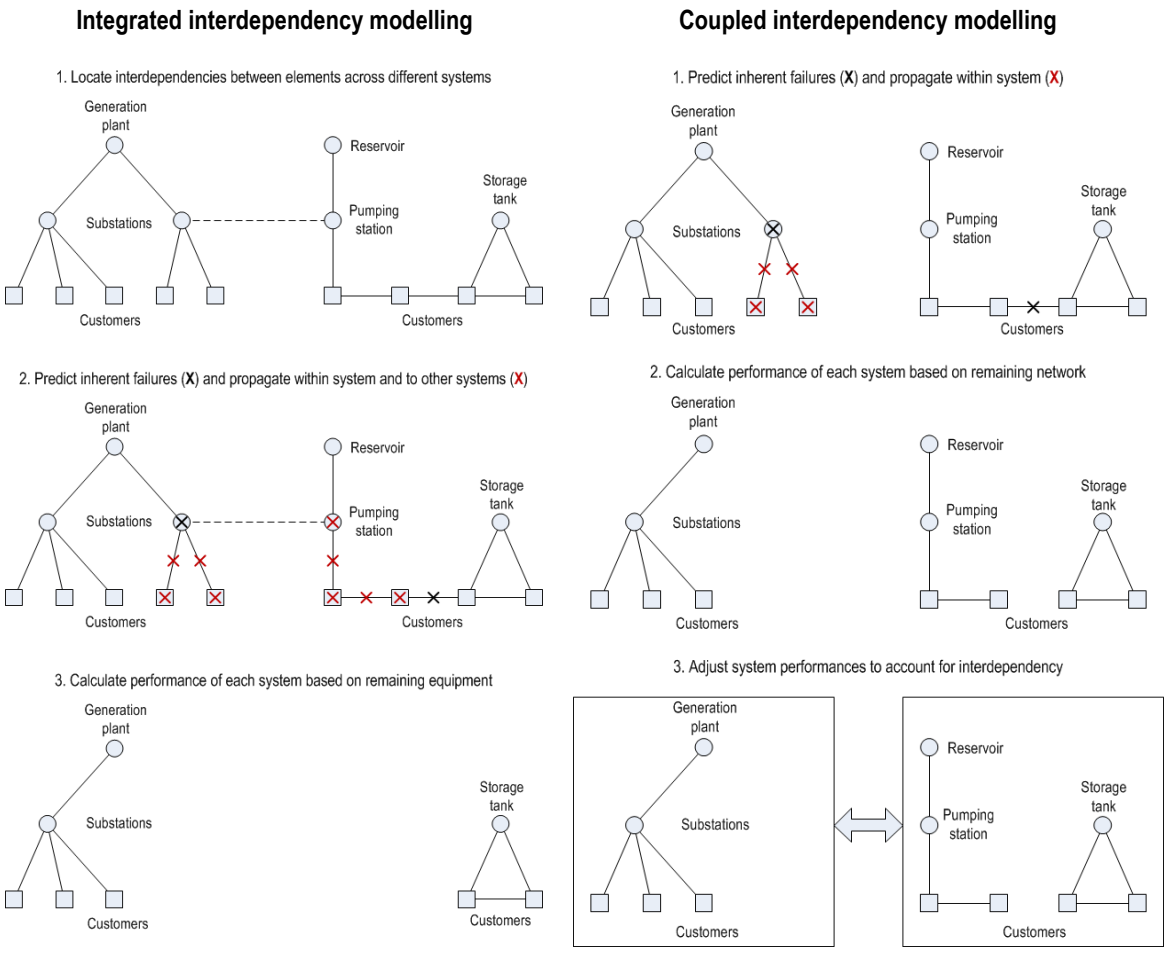

Figure 2.3 – Methodologies for high-resolution (integrated) and low-resolution (coupled) interdependency simulation illustrated with example of electric power and potable water systems

Consider an example where a substation receives power from a gas-fired power plant. If the substation fails due to earthquake damage, that failure may propagate into a pumping station, which in turn prevents a key gas processing facility from operating, since water is needed for fire safety. If that particular gas facility is critical to the operation of the generation plant, then the generation plant may fail also, so the failure of the substation has propagated back into the electric power network. Where feedback mechanisms such as this exist, the initial failures must be repeatedly propagated through each system until they are prevented from progressing any further (e.g. because the next downstream element has already failed). Depending on the

complexity of the systems and their interdependencies, this may take many cycles. Once all the failed elements are identified, they are removed from the model and the integrated system is disaggregated back into the individual lifelines, with the remaining functional elements are used to evaluate the performance of each system.

Since the integrated approach simulates interdependencies between specific elements, it is a low-level high-resolution approach (i.e. it is a micro-analysis concerned with the performance of individual components within a system). If the functional relationships between interdependent elements are deterministic and accurately defined, then this is an exact method for simulating interdependency (though uncertainty remains in hazard and fragility). The integrated approach is the most realistic representation of the true conditions but is associated with long development times, greater problems with data availability, long simulation times and non-trivial validation (Eusgeld et al., 2011). This is because as the number of systems being analysed increases and as the systems themselves become larger, so the feedback mechanisms become more complex. Another drawback of integrated models is the identification of the interdependency links. Often real connections cannot be identified and so assets across systems are assumed to be linked by geographical proximity (e.g. Dueñas-Osorio et al., 2007; Kim et al., 2009; Franchin, 2014), but the adequacy of this assumption has never been tested. Examples of integrated interdependency simulation include nodal analysis models (e.g. Zio and Sansavini, 2010; Poljansek et al., 2012; Franchin, 2014), physics-based models and agent-based simulations (e.g. Galli, 2010; Oliva et al., 2010; Eusgeld et al., 2011).

Nodal analysis models are the simplest implementation of integrated interdependency simulation. It is a connectivity-based method in which inherent failures across assets in all systems are predicted by fragility analysis and then propagated across identified interdependency links to dependent assets. The relationship between the state of the supporting asset and the state of the dependent asset can be defined deterministically (e.g. Adachi and Ellingwood, 2008; Franchin, 2014) or probabilistically. In the latter case, either a specific damage state can be assigned to the asset based on the determined propagating failure

probability so that a single event system performance can be predicted (e.g. Duenas-Osorio et al., 2007; Poljansek et al., 2012) or the system can be analysed in purely probabilistic framework (e.g. Buxton et al., 2010). The extension of nodal analysis to account for serviceability is known as physics-based analysis (Rinaldi, 2004). Agent-based models simulate the actions of 'agents' and the reactions of systems to them (Railsback and Grimm, 2011). Agent-based models are generally used to deal with complex adaptive systems, which include critical infrastructure systems since after a shock the state of assets can change both immediately and then continually over time (Rinaldi, 2004). In critical infrastructure systems, the agents are the assets and each asset is autonomous with its own performance rules to respond to a variety of system perturbations and changes to other assets in the system. The advantage of agent-based models is that they can simulate very complex relationships and their evolution over time. However, data requirements are extremely high and simulation times are very long, which affects their applicability to a Monte Carlo simulation.

Whilst for some purposes the resources to carry out such intensive integrated analysis may be justified, less resource-intensive methods are also needed for rapid assessment or where data acquisition is problematic. In the coupled approach, interdependencies between specific elements are not considered and instead interdependencies are considered at the system level. Since this is an abstract representation of the real situation, it is considered to be a high-level low-resolution approach (i.e. it is a macro-analysis that is not concerned with the performance of individual components within a system). The advantage of coupled approaches is that they are less resource-intensive than the integrated approach, but this is potentially at the cost of some precision. In coupled models inherent failures are identified and propagated only within each individual lifeline. The failed elements are removed from the model and the performance of each system is evaluated based on the remaining elements within each system. No interdependencies have been considered yet so the performance that has been evaluated is the isolated system performance. To account for interdependencies, the isolated system performances are used as inputs into a higher-level system linkage model in which the

relationship between systems is quantified, thus allowing the isolated system performance measures to be adjusted.

Examples of coupled interdependency simulation include system dynamics (e.g. Min et al., 2007; Santella et al., 2009) and input-output models (Leontief, 1986). In system dynamics models, the set of lifelines is represented by a causal loop diagram in which each lifeline is represented as a single node that is assigned an initial value determined from the isolated system performance. Interdependencies between systems are represented by links between the nodes (with separate links for each direction of dependency if the relationship between a pair of systems is bidirectional). Associated to each link is a relationship, (qualitative or quantitative) that indicates how the performance of the dependent system changes with the performance of the supporting system.

Input-output models are used in economics to assess the interdependencies between sectors of the economy but Haimes and Jiang (2001) described how this model can be applied to lifelines interdependency analysis, and this method has been adopted by other studies (e.g. Dueñas-Osorio et al., 2004; Reed et al., 2009; Setola et al., 2009; Oliva et al., 2010). It is a matrix-based methodology in which the inoperability of each system (i.e. the loss of functionality) is given by Eq. 2.46, where **x** is a vector of the final system inoperabilities, **A** is a matrix describing the interdependencies the systems and **c** is a vector describing the initial perturbation or disturbance to each system. The interdependency matrix, **A,** is populated by indices which quantify, for every pair of systems, the reduction in performance of the dependent system caused by a unit reduction in the performance of the supporting system. The indices take values of between 0 and 1. The perturbation vector, **c**, contains the inoperabilities estimated for each system before interdependencies are accounted for, i.e. loss of performance due to inherent failure only. For a model with two systems, Eq. 2.46 can be written out in full as Eq. 2.47, where a_{12} represents the effect of performance reduction in system 2 on performance in system 1, and a_{21} is the reverse. The goal of the Leontief input-output method is to solve for **x** when **A** and **c** are known, and this is achieved by rearranging Eq. 2.46 to give Eq. 2.48, where **I** is the identity matrix.

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{c} \tag{2.46}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & a_{12} \\ a_{21} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{2.47}$$

$$\mathbf{x} = \left(\mathbf{I} - \mathbf{A}\right)^{-1} \mathbf{c} \tag{2.48}$$

One advantage of this method is that it solves for all systems at the same time, whereas the systems dynamics method requires a starting point, from which it computes on a system-by-system basis. The challenge associated with both of these methods is the definition and derivation of the interdependency indices to populate the matrix **A**. This relates to the third objective considered in this thesis, interdependency quantification.

2.6.3 Interdependency quantification

The quantification of the degree or strength of interdependency between two infrastructure systems is necessary for the application of input-output models and can also be applied to increase the power of system dynamics models. However, it is also valuable as an end in itself. In summarising the dimensions associated with interdependent infrastructures, Rinaldi et al. (2001) state that the degree of coupling between two systems is one of the primary characteristics relating to the response dimension. For infrastructure operators it is useful to know how tightly coupled their system is with other infrastructure as this can help to identify vulnerabilities, threats and risks. Examples include system dependency matrices produced by the Critical Infrastructure Protection Task Force of Canada (Pederson et al., 2006) and the 'Risks and Realities' project in Canterbury, New Zealand (Lamb, 1997). These matrices identify the strength of interdependency between every pair of systems in the study region. However, in both cases the indices are qualitative, using a four-part scale from 'none' to 'high', and have been derived subjectively.

There are two major issues concerning the quantification of interdependency. Firstly defining what exactly the degree of interdependency means, and secondly developing a method for

determining it objectively. There are few studies that have attempted to address this and the primary source is the collection of studies that apply the input-output modelling method for interdependencies as described in section 2.6.2. These adopt the inoperability definition proposed by Haimes and Jiang (2001) whereby the strength of interdependency between two systems takes a value between 0 and 1 and is equal to the loss of performance in the dependent system for every unit loss of performance in the supporting system. So for example, if system 1 is dependent on system 2 with an index of 0.8, then 100% loss of performance in system 2 would result in an 80% loss of performance in system 1. The nature of this index is that it is applied to percentages and so can only be applied with performance metrics that are bounded and non-negative. Although many studies adopt this definition, only Setola et al. (2009) and Oliva et al. (2010) actually seek to derive the strengths, whilst in the other input-output studies, strengths are simply assumed in the application of a case study. However, Setola et al. (2009) derive the values subjectively by an expert judgment approach. Oliva et al. (2010) propose that the strengths can be quantified by carrying out a highly granular agent-based simulation of the systems concerned.

Rinaldi et al. (2001) qualitatively define the degree of coupling in terms of time dependence between systems such that the more tightly coupled the pair of systems, the faster disturbances propagate between them. A number of studies have built on this definition and quantified interdependency in terms of time relationships. Zimmerman and Restrepo (2009) and Galli (2010) propose measuring interdependency in terms of relative duration, which is the ratio of the outage time of the dependent system to the outage time of the supporting system. Duenas-Osorio and Kwasinski (2012) quantify coupling strengths between systems observed in the M_W 8.8 Maule, Chile earthquake by measuring cross-correlations and lag times between system restoration curves. The problem with time-based methods is that they are mostly suited to the quantification of interdependencies observed from real earthquakes. Time-based methods can be applied to simulated events, but outage and restoration times need to be predicted. Methods for doing so are few in number and increase the uncertainty associated with the quantification since

an additional modelling step is required beyond standard system operability metrics. The applicability of time-based methods is therefore restricted to systems and locations where observed data exists.

2.7 Risk metrics

The purpose of a risk metric is to distil the results from a simulation exercise into probabilistic estimates of specified outcomes in order to help decision makers understand the likelihood that their assets will negatively impacted. From the seismic risk assessment studies listed in Table 2.2, the most common risk metric is the system risk curve (e.g. Wang and O'Rourke, 2006; Shinozuka et al., 2007; Shiraki et al., 2007; Jayaram and Baker, 2010), which plots annual exceedance probability against the system loss metric. A similar metric is already used in the insurance sector. The exceedance probability (EP) curve plots estimated loss against return period (the inverse of annual exceedance probability) and is the primary metric used by insurers for risk modelling (Grossi et al., 2005). In both plots the risk is characterised annually. Therefore the loss metric must be annualised, which means that there should be only one value of the loss metric associated with each year in the event catalogue. In cases where there is more than one event in a year, an appropriate manipulation of the observed loss metrics can be applied, e.g. summation of losses, average of losses, maximum single loss. In insurance, when losses are summed, the resulting EP curve is referred to as an aggregate EP curve, and when the maximum loss value is used, the resulting EP curve is referred to as an occurrence EP curve (Grossi and Kunreuther, 2005). After the losses have been annualised, the annual exceedance probability, AEP_i, of a particular value of the loss metric, i, is given by Eq. 2.49.

$$AEP_i = \frac{N_{loss \geq i}}{N_{total}} \tag{2.49}$$

In Eq. 2.49, $N_{loss \geq i}$ is the number of times in the annualised dataset that i is equalled or exceeded, and N_{total} is the number of years in the catalogue.

2.8 Summary

The literature review has shown that there are a number of studies that have previously investigated the problem seismic risk to critical infrastructure systems. Only a few of these propose a multi-system methodological framework for conducting seismic risk assessments, and all of those that do contain weaknesses that mean that they do not fully meet the requirements for a framework for insurance-focused seismic risk assessment. The SYNER-G (Pitilakis et al., 2014a) and Hazus (FEMA, 2015) frameworks come closest, but SYNER-G does not consider financial loss and is overly complex in certain aspects, e.g. electric power network functionality, while Hazus does not adequately address interdependencies. Consequently there is a need to define an insurance-specific method for seismic risk assessment that combines applicable elements from different existing frameworks. Furthermore, whilst the existing frameworks describe methods for seismic risk assessments that make use of model components that already exist, it would be beneficial to the insurance industry to set out a procedure that would allow seismic risk assessments to be tailored according to geographical requirements. Therefore a set of model development guidelines, based on the defined seismic risk assessment method, are proposed in Chapter 3. The model development guidelines and seismic risk assessment method are demonstrated in this thesis with the case study of the electric power network and water supply system in Christchurch, New Zealand. However, the research value of this case study goes beyond just being an example application.

With respect to seismic hazard, spatial correlation models currently exist for Los Angeles, Chi-Chi (Taiwan), Japan and Europe, but none exist for Christchurch or New Zealand more generally. A spatial correlation model developed as part of the case study, using data from the Canterbury earthquake sequence, would be the first of its kind. There are a number of simplified models for predicting liquefaction triggering, and although one of these has been developed with data from the Canterbury earthquake sequence (amongst other events), its performance to other simplified models has not been compared. The model development presents an

opportunity to determine the best simplified model for predicting liquefaction triggering in Christchurch.

With respect to asset fragility, there are no existing functions that predict damage to buried cables and so a function based on data from the Canterbury earthquake sequence would be the first of its kind anywhere in the world. Whilst there are numerous repair rate functions in existence for pipes, none are specific to Christchurch and so there is an opportunity to develop a new function that is and therefore accounts for attributes of Christchurch pipes that may not be addressed by functions developed from data from other countries. Fragility functions do exist for substations, wells and pumping stations. However, these are few in number and based primarily on data from the United States and Europe. New functions developed from Canterbury earthquake sequence data will not only be the first of their kind for New Zealand, they will also add globally to the range of options for modelling fragility of these assets.

Although many studies have developed models for predicting the performance of critical infrastructure systems, the literature review has found that only four studies have previously validated models against observations from a real earthquake, with a further two examples from the literature on hurricane-induced performance issues. The validation of the case study model with observations from the Christchurch earthquake means that the model may become one of the few validated post-disaster system performance models in existence.

The study of interdependencies is a relatively recent consideration in the problem of seismic risk assessment of critical infrastructure systems. A common approach when applying integrated interdependency models is to assume assets across systems are linked by geographical proximity because physical links cannot be identified. The data provided for the case study allow both the geographical proximity and the physical link approaches to be tested and will be the first study to compare the performance of the two approaches. Finally, there are very few methods to quantify the interdependency between critical infrastructure systems. Those that do exist are either subjective, or reliant on data from observed events, or make use of methods that

are unsuitable for application by the insurance industry. There is a clear scope for developing a new method that meets the requirements of the insurance industry, with particular consideration being given to the requirement that the method should not be reliant on observed data, since such data are scarce.

3. Methodology

This chapter sets out how the three objectives of this research, defined in Chapter 1, will be met. It presents the proposed seismic risk assessment modelling framework described in objective1, which includes methodologies for model development and model application. It goes on to present an introduction to the case study in Christchurch, New Zealand, to which the framework will be applied to satisfy objective 2 and finally it describes the methodology for interdependency quantification that will be applied to the results of the case study seismic risk assessment to fulfil objective 3.

3.1 Modelling framework

The review of existing studies presented in Chapter 2 shows that methodologies adopted for seismic risk assessment of lifelines in existing studies are generally similar. They broadly follow the equivalent procedure already implemented within catastrophe models for building portfolio risk assessment with inventory exposure as the input data and three modules: hazard, vulnerability and loss. Within each of these modules, the specific models and functions applied at each step vary from study to study, and it is possible to frame the different approaches in terms of a methodological hierarchy, as summarised in Table 3.1, where 'Low' refers to the least detailed analysis option and 'High' refers to the most detailed analysis option.

Whilst one would ideally adopt the methods at the highest hierarchy level, this is not always feasible due to resource and data availability constraints. To meet the needs and constraints of the insurance and catastrophe modelling, an industry-specific methodology is proposed in which the basic modules of seismic risk assessment are sub-divided into eight steps. The following sections describe each of the steps in terms of both model development and model application. The overall procedure for model development is summarised diagrammatically in Figure 3.1 and the method for model application is summarised diagrammatically in Figure 3.2. The proposed method pulls together elements from different studies reviewed in Chapter 2 and can

be considered to be the best practice for the seismic risk assessment of lifelines in catastrophe modelling. It is a general methodology that can be applied to any case study and will be the basis of the model development in Chapter 4.

Table 3.1 – Methodological hierarchy in lifelines risk analysis

Hierarchy Level	Exposure	Hazard	Vulnerability derivation	System Performance	Interdependency
Low	Key facilities only	None, 'what-if' scenarios	Expert judgment	No systemic analysis	None
Medium	Up to distribution level or equivalent	Single scenario earthquake	Empirical	Connectivity analysis	Coupled model
High	Detailed to property level	Monte Carlo simulation	Analytical	Serviceability analysis	Integrated model

3.1.1 Step 1 – Exposure

The purpose of the exposure step is to identify and classify the assets within each system being analysed. This includes their location, their function, their physical attributes and the adjacent assets to which they are connected, both within the system and if applicable within other systems too. Based on the review of existing practice in Chapter 2 and Appendix B, Table 3.2 lists the main asset types in each system whose existence and attributes should be recorded in an exposure database.

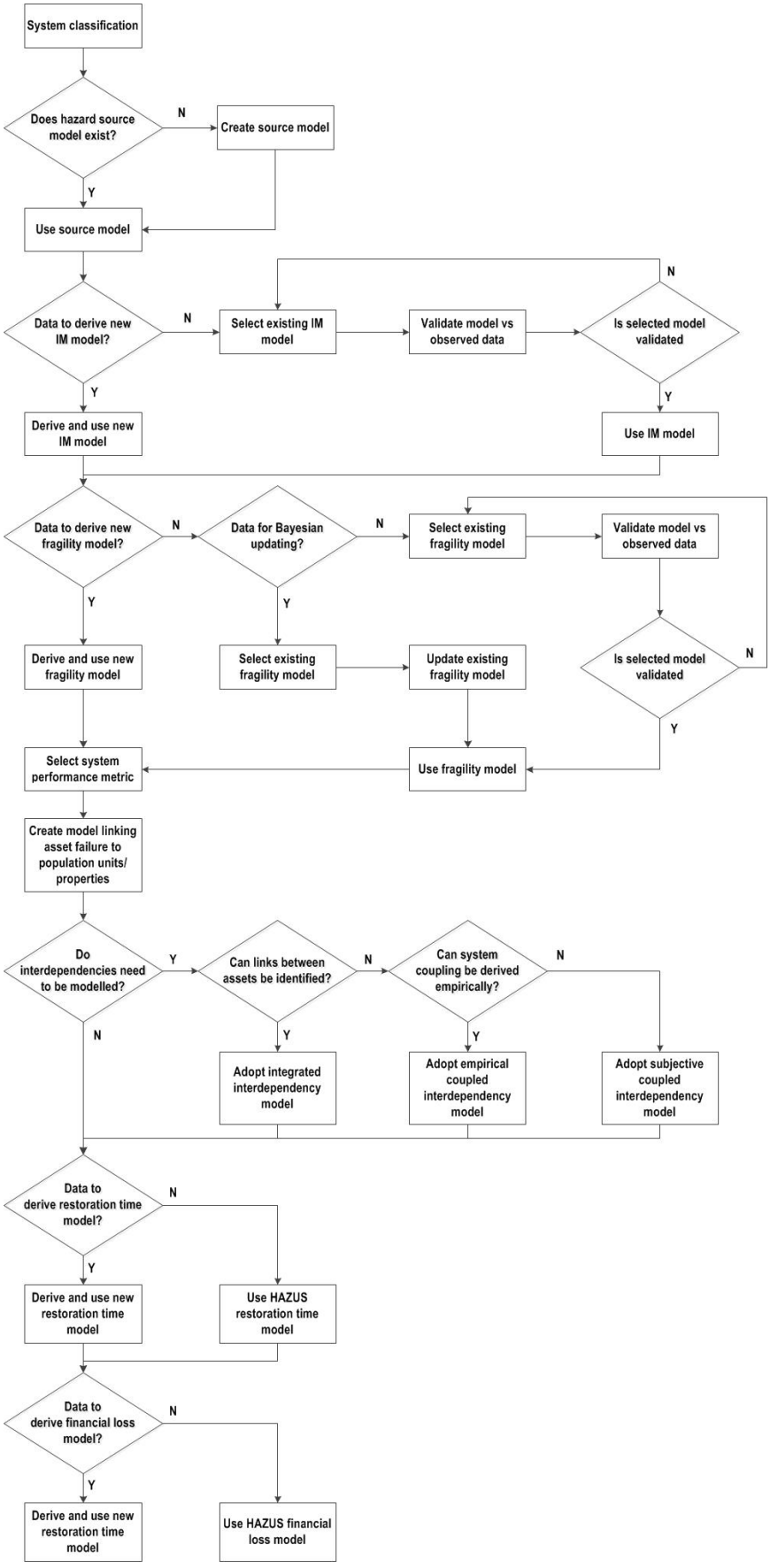

Figure 3.1 – Seismic risk assessment model development guidelines summarised as a process flowchart

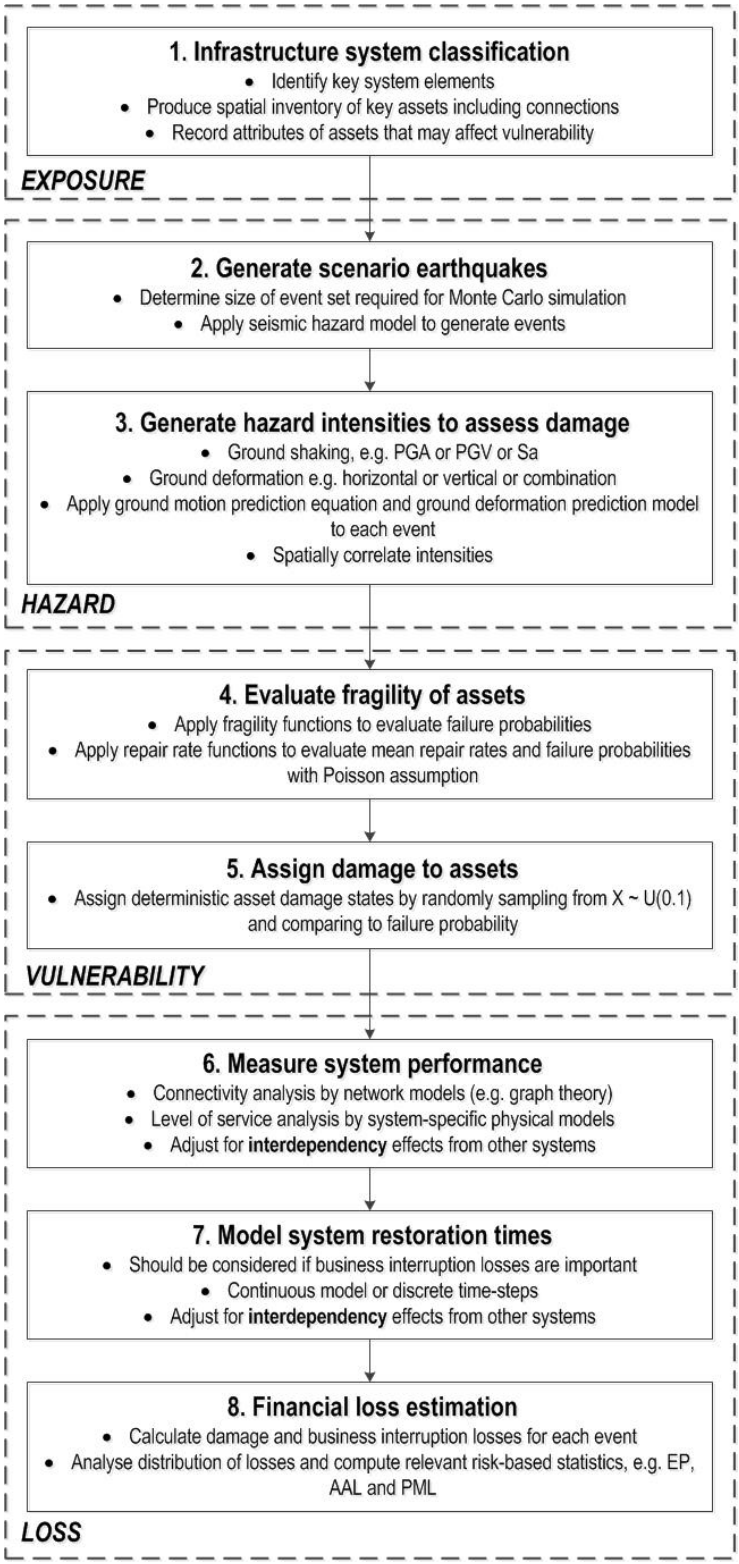

Figure 3.2 – Best practice generalised procedure for seismic risk assessment of lifelines systems based on existing studies

Table 3.2 – Primary asset types to be considered in seismic risk assessment of lifelines systems and relevant attributes for documentation

Infrastructure system	**Components**	**Component attributes**
Electric power	Generation plants	Capacity, seismic design level
	Substations	Voltage, seismic design level
	Cables	Material, size
Potable water	Wells	Seismic design level
	Water treatment plants	Capacity, seismic design level
	Pumping stations	Capacity, seismic design level
	Storage tanks	Elevation, material, geometry, quantity of contents, seismic design level
	Pipes	Material, joint type, age, diameter
Waste water	Lift stations	Capacity, seismic design level
	Treatment plants	Capacity, seismic design level
	Pipes	Material, joint type, age, diameter
Natural gas	Pipelines	Material, joint type, age, diameter
	Compressor stations	Capacity, seismic design level
Fuel	Refineries	Capacity, seismic design level
	Pumping stations	Capacity, seismic design level
	Storage tanks	Elevation, material, geometry, quantity of contents, seismic design level
	Pipelines	Material, joint type, age, diameter
Telecommunications	Central offices	Seismic design level
	Cables	Material, size
Highways	Roadways	Importance level
	Bridges	Structural system, material, age, geometry, seismic design level
	Tunnels	Construction method, geometry, local geology
	Embankments	Height, soil type

3.1.2 Step 2 – Event generation

For infrastructure systems it is necessary to analyse performance based on scenario events rather than probabilistic hazard. To establish a probabilistic view of risk it is also necessary to carry out a Monte Carlo simulation of events, and to account for uncertainty the number of events in the simulation must be sufficiently large to minimise variance in the output statistics. Rougier (2013) states that the confidence interval around a Monte Carlo simulation statistic can be estimated for a specified number of simulations by applying the Dvoretzky-Kiefer-Wolfowitz Inequality. With this method the 95% confidence interval is estimated to be ±4.3% relative to the mean value for 1,000 simulations; ±1.9% for 5,000 simulations; and ±1.4% for 10,000 simulations. Therefore the number of simulations can be chosen to reflect the requirements of the analyst to minimise variance. In many locations an earthquake source model may already exist from which events can be stochastically generated based on the assumption that the time interval between events follows a Poisson distribution. If this is not the case, an a bespoke source model can be developed by following the procedure of Coburn and Spence (2002), where data from historic and instrumental catalogues can be geospatially analysed to identify sources and corresponding Gutenberg-Richter relations.

3.1.3 Step 3 – Hazard intensities

The first part of generating hazard intensities is to identify which IMs need to be predicted, which depends in part on the fragility functions selected. For ground shaking induced damage, the significant majority of fragility functions for lifelines assets are expressed with respect to PGA (above ground assets) and PGV (buried assets) and so prediction of these two IMs should be sufficient in most cases. The only other ground shaking IM that is occasionally used is SA and so if an SA-based function is selected in step 4, then the hazard intensity generation will need to be revisited. As documented by Douglas (2015), there are numerous existing GMPEs for predicting ground shaking intensity and so it is not considered to be necessary to develop

new empirical relationships even if data are available. It is however essential to select one that is relevant to the seismo-tectonic environment being analysed.

A spatial correlation model must also be applied to correlate the GMPE residuals. In some locations a spatial correlation model may already exist but otherwise, as discussed in Chapter 2, there is a range of methods for developing bespoke empirical correlation models given a sufficient quantity of local strong motion data (e.g. Boore et al., 2003; Wang and Takada; 2005; Goda and Hong, 2008; Goda and Atkinson, 2009; Jayaram and Baker, 2010). No one method is considered to be superior over the others, and so it is suggested that as many of these are tested as possible to ascertain the best spatial correlation model for a specific case study. If no data are available to create a new correlation model, then a model should be selected from the existing literature that has been developed for a location that matches as closely as possible the geological environment of the case study. The correlation model can be applied using the method of Weatherill et al. (2014).

For liquefaction-induced ground deformation it is necessary to predict first whether liquefaction is likely to occur. This requires analysis of liquefaction susceptibility maps to determine if liquefaction can occur and application of a liquefaction triggering method to classify whether liquefaction does occur. Liquefaction susceptibility maps may already exist in some locations. If not, then the FEMA (2015) method for liquefaction prediction includes a table classifying liquefaction susceptibility for different types of geological conditions, which can be compared to geological maps of the case study area to determine the closest match. For liquefaction triggering, if sufficient geotechnical data are available then the detailed method of Seed and Idriss (1967) should be applied. However, in an insurance context it is more likely that a simplified method will be required. Andrus and Stokoe (2000), Hazus (FEMA, 2015) and Zhu et al. (2015) propose simplified methods for predicting likelihood of liquefaction. If empirical data are available, it is suggested that all three are tested to ascertain the best liquefaction triggering model for a specific case study.

If liquefaction is predicted to occur, then its scale in both the horizontal ($PGDf_H$, lateral spread) and vertical ($PGDf_V$, differential settlement) directions should be predicted. Only Rauch and Martin (2000) and Hazus (FEMA, 2015) propose simplified methods appropriate for insurance application. The model by Rauch and Martin (2000) only predicts horizontal movements and so for consistency between vertical and horizontal movements, it is recommended to use the Hazus model. The lack of existing models means that there is no accepted functional form for a PGDf prediction equation (as there is for GMPEs) and there is no guidance on which predictor variables should be included. Development of a new PGDf prediction equation would require specialist scientific expertise and it is assumed to be unlikely that this will be available to an insurer. It is therefore not recommended to attempt to develop new empirical models even if observed PGDf data are available.

PGDf induced by surface fault rupture can be predicted by the equation of Wells and Coppersmith (1994). However this is only necessary in assessments where there are seismic sources directly under the study area. If sufficient data on geological conditions and slope angles are available, then the method of Saygili and Rathje (2008) can be applied to predict PGDf due to landslides. However, it is anticipated that this will not be feasible in the majority of insurance studies.

3.1.4 Step 4 – Fragility evaluation

Asset damage can be described in physical terms or functional terms, with the former being more important for estimating repair costs and the latter more important for estimating system performance and business interruption effects. The relevant fragility function is applied to each asset to predict the probability of being in a particular physical damage state based on the IM. If system performance is modelled, it is necessary to map functional states. Partial functionality can be considered by assuming a certain level of functionality given a particular damage state but there is little literature on this and partial functionality is only of value if serviceability system performance analysis is being applied. For pipes, the convention is for damage to be

measured on a continuous scale as a repair rate (i.e. repairs per kilometre). Individual pipes are grouped by typology and then within each group, sub-groups are created of pipes that are subjected to the same range of IM values. For each sub-group the mean repair rate is calculated from a repair rate function. It is assumed that the distance interval between repairs in a group follows a Poisson distribution with a rate equal to the calculated mean repair rate. This assumption allows the probability that an individual pipe has failed to be calculated (i.e. the probability that there is at least one repair on the pipe).

As listed in Appendix C, there are many existing fragility functions available for lifelines assets. However, it may be beneficial to create new bespoke fragility functions to reflect the specific physical properties of the assets being studied. The most viable option for creating new fragility functions is to use empirical damage data and follow the guidelines of the Global Earthquake Model (GEM) (Rossetto et al., 2014). If there is only a small empirical dataset, which is deemed insufficient for creating a new fragility function (e.g. a past event that resulted in no damage), then the Bayesian updating method described in Chapter 2 can be applied to an existing fragility function to make it more relevant to the case study. If no empirical dataset exists then a fragility function can be selected from the existing literature that has been created for an asset whose physical properties most closely match the asset being studied.

3.1.5 Step 5 – Damage assignment

For system performance analysis, the failure probability of an asset is converted to a deterministic functional state by sampling a random variable that is uniformly distributed between 0 and 1, and comparing the sampled value to the failure probability. If the sampled value is less than or equal to the failure probability, the asset is designated as having failed. Otherwise it is designated as operational. In cases where liquefaction is being modelled, the functional state of each asset must be assessed in response to both ground shaking and liquefaction effects.

3.1.6 Step 6 – System performance

System performance can be modelled in terms of connectivity or level of service (serviceability). The latter is an extension of connectivity analysis and is more informative, but requires detailed technical knowledge of the physical processes involved in the flow of utility, and is therefore less likely to be of relevance for insurance application. Connectivity performance should be measured in terms of number of people affected or number of properties affected rather than simply by area. The latter can be misleading in loss calculations since it does not account for population or building density. To measure performance in terms of population, the spatial distribution of people in the study location must be known. This can be obtained from Census and administrative boundary data if this is accessible. The administrative boundary data can be used to divide the study area into smaller units and population data for each unit can be attributed from Census information. A method must then be adopted for determining whether each unit is connected to the system. In some cases infrastructure operators may publish zone maps that identify which specific assets areas are connected to, and so the connectivity of a unit is related to the operational state(s) of the specified asset(s). If these are not available then some assumption must be made, e.g. the connectivity of a unit is related to the operational state of the closest asset. To measure performance in terms of property requires a map identifying individual properties. For small study areas this may be feasible from satellite imagery such as Google Earth, but for larger study areas this may require access to proprietary databases. For loss estimation it may be useful to adopt an equivalent performance metric that measures the economic value in a study area that is affected by loss of connectivity. This requires geospatial data on the distribution of economic activity in the study area.

If multiple interdependent systems are being studied then this must be accounted for, ideally with an integrated interdependency model if there is sufficient data to do this. Otherwise, a simplified coupled model could be adopted. Nodal analysis is the simplest of the integrated interdependency simulation methods to apply and therefore the most useful for insurance purposes. It requires the links between assets within each system to be defined and this may be

achieved with information already gathered in the exposure database. As with the population maps for connectivity performance, if operators publish zone maps, then it may be possible to identify the genuine links between assets. Otherwise, an assumption such as geographical proximity could also be used. There may be situations where integrated interdependency simulation becomes extremely complex (e.g. when there are many systems with multiple feedback relationships), in which case a coupled approach such as system dynamics or Leontief input-output modelling may be more appropriate. This requires definition of an index to quantify the strength of the link between systems. Ideally this should be derived from empirical data (not necessarily only from earthquakes), but this type of information has not been well-recorded in the past. Alternatively, an infrastructure operator may be able to provide a subjective assessment of the dependency of their system on others.

3.1.7 Step 7 – System restoration

Modelling of restoration times is important for modelling business interruption losses. For insurance purposes, discrete time-step modelling is more appropriate than continuous modelling. Even when an asset is repaired after an earthquake it is still subject to the effects of interdependencies and so this must be re-modelled at each time-step. If a system restoration model is applied, this allows more detailed system performance metrics to be used that account for the outage time of assets, such as the measurement of outages in terms of people minutes. There are few existing models to for system restoration (e.g. Cagnan et al., 2006; FEMA, 2015). If sufficient empirical datasets exist then ideally new restoration functions should be developed to allow at least for a time-step analysis.

3.1.8 Step 8 – Financial loss estimation

Losses associated with physical damage and business interruption are calculated for each event in the stochastic catalogue and the relevant statistical analysis is undertaken according to the requirements of the insurer. This should at a minimum typically include exceedance probability curves, average annual loss and probable maximum loss. To convert physical damage to losses

it is necessary to predict the mean damage ratio, which is the repair cost as a proportion of the total replacement value of the asset. This can be done empirically by deriving models from historic data that link damage ratio to physical damage state or directly to IM. Alternatively, FEMA (2015) have published a comprehensive list of tables linking physical damage states to damage ratios for a wide range of assets. For assessing business interruption losses, it is necessary to use a model that links losses to restoration time. If sufficient datasets are available this can be done empirically, or the method proposed by FEMA (2015) can be used.

3.2 Scope of case study application of framework

The proposed model development framework is applied to a case study of the electric power and water supply systems in Christchurch, New Zealand. A seismic risk assessment model will be developed for these two systems based on empirical data from the Canterbury earthquake sequence that affected Christchurch in 2010 and 2011. The model will be validated against observations from the Canterbury earthquake sequence and will then be applied to evaluate a future projection of risk for these two systems. This section presents contextual information about the Canterbury earthquake sequence and the engineering performance in those events. It goes on to outline the specific elements of the model development guidelines that will be implemented in the case study, the validation process for the model and the diagnostics used to evaluate risk when the final validated model is applied for future projection.

3.2.1 The Canterbury earthquake sequence

Christchurch is located in the Canterbury region on the South Island of New Zealand and is the second largest city in the country, with a population of around 350,000. The main economic activity in Christchurch is agri-business, in particular the processing of agricultural produce from the surrounding Canterbury region, but there is also a strong manufacturing base in the city. This focus on primary and secondary sectors means that the economic vitality of Christchurch is strongly linked to the ability of the electric power and water supply systems to meet commercial demands. The Canterbury earthquake sequence occurred between 2010 and

2011 and generated in excess of 10,000 individual seismic events, mostly of moderate magnitude. However, their proximity to urban centres meant they resulted in significant damage to buildings and infrastructure, particularly in Christchurch. For empirical model development and model validation, this thesis focuses on the two most damaging events within the sequence: the initial M_W 7.1 main shock on 4th September 2010 (the Darfield earthquake) and the M_W 6.2 aftershock on 22nd February 2011 (the Christchurch earthquake). A context map of New Zealand and the locations of the epicentres of the two earthquakes in relation to the city of Christchurch are shown in Figure 3.3 and Figure 3.4 respectively. The key features of these two events that are relevant to this research are summarised here, but a more detailed treatment of the ground motion and seismic source aspects of the sequence can be found in Bradley et al. (2014).

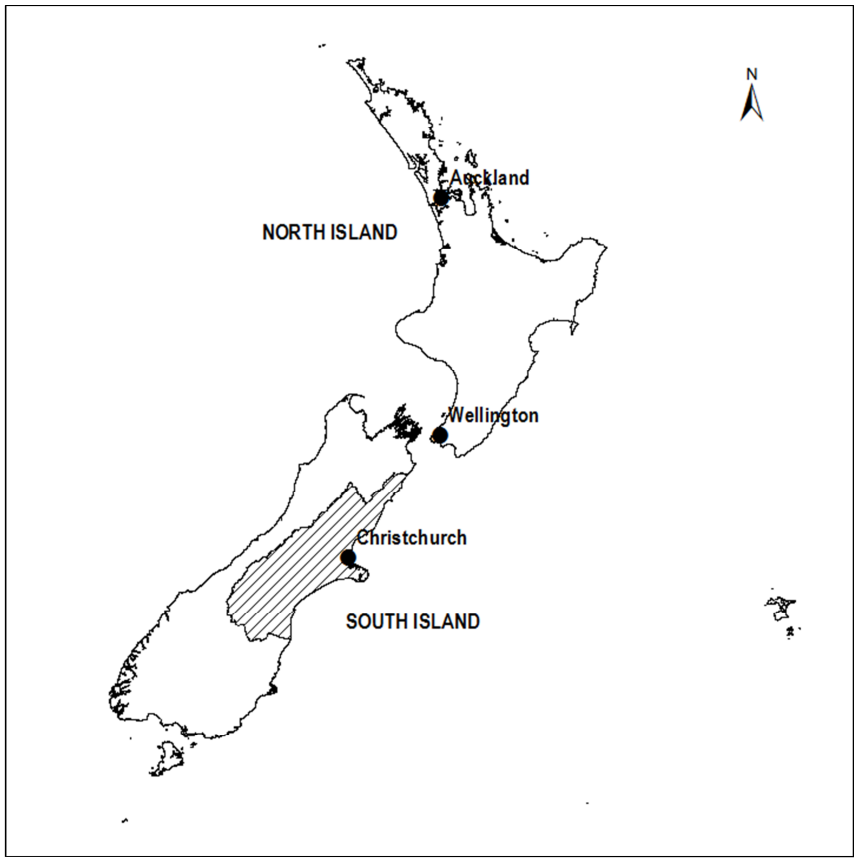

Figure 3.3 – Map of New Zealand showing highlighting Christchurch, the Canterbury region (shaded grey) and other major cities

The initial shock was the largest event in the sequence with its epicentre near the town of Darfield, approximately 40 km west of Christchurch and at a depth of 8 km. It was the first large earthquake to strike close to an urban centre in New Zealand since the M_W 7.4 Hawke's Bay earthquake of 1931. Four rupture sources were identified but the most energetic of these was the strike-slip rupture of the previously unrecognised Greendale Fault. Accelerograms were recorded from 130 sites across South Island, ten of which had PGAs in the range 0.3 g to 0.82 g. 18 records showed PGVs exceeding 0.5 m/s, with three of them exceeding 1 m/s. It resulted in no fatalities and only two serious injuries (Wood et al., 2010). In part this was due to the good performance of buildings and also because the earthquake occurred in the middle of the night, so few people were on the streets. The majority of building damage was observed amongst historic and unreinforced masonry (URM) buildings. Although only two URM buildings completely collapsed, many suffered irreparable damage and had to be demolished (Dizhur et al., 2010).

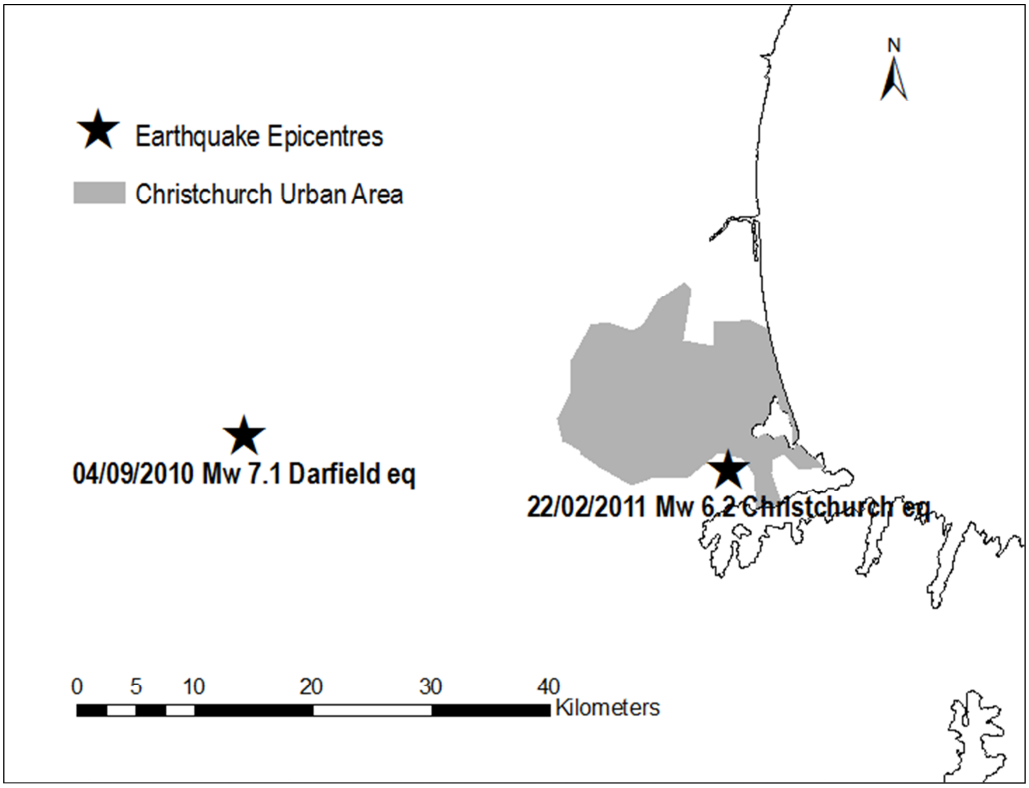

Figure 3.4 – Location of epicentres of the Darfield and Christchurch earthquakes in relation to the Christchurch urban area and central business district

Modern and seismically retrofitted URM buildings generally performed well. Some minor damage occurred to substations and buried cables in the electric power system, but most customers had power restored by the end of the day. The water supply system was almost fully restored within five days. However, the waste water systems was more severely affected, with some customers still relying on temporary sewage facilities after two months. Some minor disruption occurred to the telecommunications systems and no damage was observed in the gas system (Wood et al., 2010).

The 22nd February aftershock was the most damaging event in the sequence. It was caused by a strike-slip rupture on a previously unrecognised fault, with an epicentre 10 km to the southeast of the central business district (CBD) and a depth of 5-6 km. The shallow depth and proximity created very high ground motions in the city itself. The PGA at the Christchurch Botanic Gardens in the CBD was 0.5 g in the horizontal direction and the highest PGA recorded was 1.41 g (Bradley and Cubrinovski, 2011). This event caused 185 fatalities and 8,600 injuries (Giovinazzi et al., 2011) and occurred while the Canterbury region was still recovering from the Darfield earthquake, with many structures suffering from compounded damage. The majority of damage occurred to URM buildings but in contrast to the Darfield earthquake, significantly more damage was observed in reinforced concrete (RC) buildings. 16% of RC buildings in the CBD collapsed (Kam et al., 2011) and 135 of the fatalities occurred across two RC building collapses. Large parts of the city suffered power outages immediately after the earthquake, primarily due to damage to buried cables but also some minor damage to substations. The water supply system suffered extensive damage to pipes and it is estimated that 80% of the city population lost mains connection immediately after the earthquake (Ministry of Health, 2012). The waste water system also experienced extensive damage to pipes and the main waste water treatment plant in the city suffered severe damage resulting in a major capacity reduction. The city relied heavily on temporary sewage facilities, in some cases for over a year afterwards and there were potential health risks due to cross-contamination between the two water systems. There were numerous road closures and five of the six bridges crossing the River Avon were

closed. The gas system performed extremely well with no observed damage, although fire safety concerns around the CBD resulted in services being manually shut off as a precaution (Giovinazzi et al., 2011).

Both events, but in particular the Christchurch earthquake, caused unprecedented levels of liquefaction throughout the southern and eastern suburbs of Christchurch alongside the Avon River. The liquefaction resulted in settlement, lateral spreading, sand boils, and a large quantity of ejected silt mud and water ponding on the ground surface. This was a consequence of the alluvial deposits and high water table (between 0-3 m) that characterize the ground conditions in these suburbs (Giovinazzi et al., 2011; Yamada et al., 2011). In the Darfield earthquake, 15,000 residential dwellings were affected by liquefaction, of which at least 1,000 were completely destroyed (King et al., 2014). Initially geotechnical engineering solutions were proposed to mitigate liquefaction risk from future earthquakes and ensure that affected suburbs remained habitable. However this proposal was abandoned after the Christchurch earthquake. Fresh liquefaction caused land close to the River Avon to sink further, effectively raising the level of the river by up to 0.5 m and generating unacceptable flood risk for nearby properties (Hughes et al., 2015). The Canterbury Earthquake Recovery Authority (CERA) initiated a policy to designate all residential land in Christchurch according to its liquefaction risk. Some land was so badly damaged that it was deemed uneconomic to repair and was designated as being in a 'Red Zone'. Over 7,000 properties have been abandoned because of a 'Red Zone' designation (King et al., 2014). After the Christchurch earthquake, differential movement of foundations affected many commercial buildings in the CBD. The damage was so widespread and the risk of tall building collapse so great, that the army was called in to man a temporary exclusion zone around the city (Giovinazzi et al., 2011). The cordon gradually reduced in size, but was not lifted until June 2013 and roads in the CBD did not completely re-open until July 2014. Over 1,500 commercial properties in the CBD had to be demolished (Saunders et al., 2014) and because of the cordon, all 6,000 businesses in the CBD, even those whose property was undamaged, were forced to relocate (Stevenson et al., 2012).

New Zealand operates a compulsory government scheme for residential earthquake insurance, managed by the Earthquake Commission (EQC). They received 150,000 claims for residential building damage after the Darfield earthquake and a further 40,000 for land damage. After the Christchurch earthquake they received 140,000 claims for residential building damage and a further 60,000 for land damage (King et al., 2014). Swiss Re estimates total losses of 6 billion USD from the Darfield earthquake and 15 billion USD from the Christchurch earthquake. Analysis by King et al. (2014) shows that compared to other major loss-causing earthquakes in the period 2009-2011, the two earthquakes caused relatively high losses as a percentage of gross domestic product (15%), However 80% of the these losses have been covered by the insurance sector, considerably more than for the next most insured (27% for the 27th February 2011 earthquake in Chile).

3.2.2 Model development

The research focuses on engineering issues and so the case study model will not include step 8 in Figure 3.2 relating to financial losses. Also step 7, relating to restoration times, is not considered due to insufficient data being available for model development and validation. The final loss metric for the model will therefore be a measure of system performance. Since serviceability analysis is of little value to the insurance sector, system performance will be measured only in terms of connectivity. Because of the need for the model to be validated, the specific nature of the performance metric adopted in the model will depend on what performance-related information is known or measurable from the validation events.

3.2.3 Model validation

In order to conduct the model validation it is necessary for the model to evaluate a system performance metric for which the value is known for the validation event(s). No performance information is known for either system in relation to the Darfield earthquake. For the Christchurch earthquake, it is known that 80% of the population lost connection to the water supply system (Ministry of Health, 2012). For the electric power, Orion are the local

distribution network operator, and have measured system performance in the Christchurch earthquake terms of customer outage minutes, which is the sum product of the disconnection times of each individual customer. However, the time aspect of system performance is not being considered in this study. As an alternative, the spatial extent of customers affected in the immediate aftermath of the Christchurch earthquake is known from an outage map provided by Orion. Orion have not provided a dataset that would enable the number of customers (i.e. households and businesses) in a specified area to be determined, but using data from the 2006 Census (Statistics New Zealand, 2014a), it is possible to determine the residential population in a specified area and so the number of affected residents, as distinct from customers, can be used to measure of system performance. Cross-referencing the outage map with the spatial population data finds that 71% of the population became disconnected from the electric power network in the Christchurch earthquake. There is no known publicly available dataset that describes the spatial distribution of economic activity within Christchurch. Furthermore, there are no data that describe the economic impact of the two earthquakes that is specifically attributable to electric power network or water supply system outage, which would be required for validation.

Therefore it is proposed that the model should use the percentage of the total population in the study area who become disconnected as the performance metric for both systems. If the percentage is expressed as a proportion, which takes a value between 0 and 1, then the metric is equivalent to the system inoperability metric proposed by Haimes and Jiang (2001) for the Leontief input-output interdependency methodology. The values of the performance metrics calculated by the model will be validated against the two observations from the Christchurch earthquake – the proportion of total population disconnected from the water supply system (80% or 0.8) and the proportion of total population disconnected from the electric power network (71% or 0.71)

To validate the model, it will be run with known input parameters (magnitude and location) for the Christchurch earthquake. However, the model will include various sources of uncertainty,

relating to ground motion prediction, prediction of liquefaction-induced permanent ground deformations and the application of fragility and repair rate functions to assets, all of which require an element of random sampling. To account for this, it is proposed is run the model 1,000 times, with random variables re-sampled each time.

3.2.4 Model application

It is proposed to apply the final seismic risk assessment model with a Monte Carlo simulation of 10,000 years' worth of events. Risk will then be evaluated for both systems at three scales: system-level, asset-level and neighbourhood-level. System-level risk will be evaluated by plotting the distribution of predicted system performances from all events and by plotting an occurrence exceedance probability curve to identify the expected system performance for specified return periods. Asset-level risk will be evaluated by counting the number of times each asset is predicted to fail in the event set. This information can be used to identify the most vulnerable assets, which may be prioritised for upgrading. Neighbourhood-level risk will be evaluated by sub-dividing the study area into smaller geographical units known as meshblocks, which are used statistical reporting in New Zealand (Statistics New Zealand, 2014b), and counting the number of times in the event set that each meshbock loses connection to the system. This information can be used to identify the most at-risk neighbourhoods and inform strategies for improving the reliability of services to these locations.

3.3 Interdependency quantification

Interdependency quantification is useful for risk management purposes and it is proposed that this can be achieved by making use of seismic risk assessment models, which are already widely used within the insurance sector. The rationale is to use the results from a seismic risk assessment to generate an objective quantitative measure of how tightly coupled two infrastructure systems are. The concept of quantifying interdependency as a single index is already proposed in the Leontief input-output method (Haimes and Jiang, 2001), although in that study the index or indices are already known and the objective is to estimate the final

system performances based on the indices and the initial pre-interdependency system performance. However, when using the method proposed here in section 3.1, seismic risk assessment models for critical infrastructure systems, estimate both the final and initial performances of each system. Therefore, if the estimated system performances are expressed as proportions (between 0 and 1), the problem defined by Haimes and Jiang (2001) can be restructured to make the objective to estimate the interdependency indices that populate the matrix **A** in Eq. 2.46.

One of the benefits of the Leontief input-output method when solving for final system performance is that all system performances can be estimated at the same time rather than iteratively. However, Eq. 2.46 cannot be rearranged to make **A** the subject and so the method proposed here requires the indices to be estimated one at a time for each uni-directional relationship. If there are two infrastructure systems where system 2 is dependent on system 1, then the final interdependent performance of system 2, SP_2 can be calculated from Eq. 3.1, where SP_1 is the performance of system 1, SP'_2 is the performance of system 2 before interdependencies are accounted for (i.e. the independent performance) and I_{21} is a constant. For the purposes of this research a new term, 'interdependency response' is proposed. In this example it is equal to the difference between the independent and interdependent performances of system 2 (i.e. $SP'_2 - SP_2$) and therefore also to the term $I_{21}SP_1$ as shown in Eq. 3.2.

$$SP_2 = I_{21}SP_1 + SP_2^{'} \tag{3.1}$$

$$Interdependency\ Response = I_{21}SP_1 = SP_2 - SP_2^{'} \tag{3.2}$$

The formulation in Eq. 3.1 implies that whilst the interdependency response will vary from event to event depending on the performance of system 1, it is constrained by the constant I_{21}. When $SP_1 = 1$, the interdependency response in system 2 is equal to I_{21}, so I_{21} can therefore be interpreted as a measure of the expected interdependency response in system 2 when there is a total failure of system 1. This is an objective metric for quantifying the capability of a system to

withstand total failure in a supporting system (and therefore how interdependent those two systems are), with the advantage that it is a single global value that is independent of any specific event and is therefore useful for high-level risk management and emergency planning purposes. From here on, the term I_{21} is referred to as the 'interdependency index', representing the potential effect of the performance of system 2 on the performance of system 1.

Defining and quantifying the interdependency according to the Leontief input-output method has the additional benefit that once the interdependency index is calculated, it can be used as an alternative to the nodal analysis method for simulating interdependencies more quickly in subsequent seismic risk assessments. However, there is one drawback with using the formulation of Eq. 3.1 for this purpose. The Leontief input-output method was developed initially to model economic sector productivity, which numerically is unbounded in the positive direction, i.e. its only constraint is non-negativity. For infrastructure system performance however, metrics relating to the number of people or properties affected have an upper bound. If the metric reports as a total number, then the upper bound is the total population or the total number of buildings served. If the metric reports a proportion, the upper bound is 1. However, if one applies Eq. 3.1 to estimate SP_2, there is nothing to prevent SP_2 being estimated as greater than the upper bound. To overcome this a new modified formulation of the Leontief input-output formulation is proposed here as Eq. 3.3, where system performance is reported as a proportion with respect to the upper bound. Within Eq. 3.3 the bracketed term is equivalent to the proportion of customers who remain connected after inherent physical asset damage is accounted for but before interdependencies are considered and ensures that SP_2 is less than 1. This is more easily demonstrated by rearranging Eq. 3.3 to give a new mathematical definition for the interdependency response, in proportional terms, as shown in Eq. 3.4. In this formulation, the definition of I_{21} remains the same, i.e. is the expected interdependency response in system 2 when there is total failure in system 1. Since the interdependency response cannot exceed 1, I_{21} cannot exceed 1 either. Therefore $I_{21}SP_1$ cannot exceed 1, which in turn ensures that the maximum possible value of SP_2 in Eq. 3.3 is 1.

$$SP_2 = I_{21} SP_1 \left(1 - SP_2^{'}\right) + SP_2^{'} \tag{3.3}$$

$$I_{21} SP_1 = \frac{SP_2 - SP_2^{'}}{1 - SP_2^{'}} \tag{3.4}$$

Both Eq.'s 3.2 and 3.4 imply that the index I_{21} can be estimated as the gradient of a linear plot of the interdependency response against the performance of system 1. The validity of this assumption will be tested in the case study application by using the outputs from the system performance and interdependency simulation analysis to create plots of interdependency response against electric power network performance using both the original and adapted Leontief methods. If the equations are valid, one would expect to observe a linear trend in the data. Also, since there is no intercept term in either equation, one would expect a best-fit line through the data to pass very close to the origin (this makes physical sense since the interdependency response should be zero when there is no loss of performance in system 1). The slope of this best-fit line is the interdependency index, I_{21}, and this method is referred to as the 'linear slope method'. If the best fit line passes through or close to the origin, then the slope is approximately equal to the expected performance when SP_1 is equal to 1 (i.e. total failure in system 1). This approximation facilitates an alternative method for estimating the interdependency index. If within the plotted data there are multiple events where SP_1 is equal to 1, then the interdependency index can be approximated as the mean of the interdependency responses just for those events. This is referred to as the 'extreme event method' and is useful in situations where the performance of one of the systems cannot be modelled explicitly, e.g. due to unavailability of data or resources. Both methods are applied to the Christchurch case study to estimate $I_{WSS|EPN}$, the interdependency index for the effect of the electric power network on the water supply system. For this specific case study there is little benefit in the application of the extreme event method since both systems can be modelled explicitly, but the latter is applied to compare its results to those obtained by the slope method.

Two sensitivity tests are proposed to investigate the effect of certain system properties on the index. The first sensitivity evaluates the effect on the interdependency index of the fragility of assets in the water system. This is done by repeating the fragility, system performance and interdependency analyses for all events whilst varying the parameters of the pipe repair rate functions and parameters of the fragility functions used for wells and pumping stations. The second sensitivity test evaluates the effect that the quantity of back-up power supplies has on the index by repeating the interdependency analysis for all earthquake events whilst randomly varying the number of wells and pumping stations that have access to back-up power.

4. Model Development & Validation

This chapter describes in detail the development of the seismic risk assessment model for the electric power network and water supply system in Christchurch. Individual sections are presented for each of the model components identified in Chapter 3 (see Figure 3.2), with the exception of restoration time and financial loss modelling, which are not addressed in this research. Finally, the outcome of the model validation with respect to the system performance observations from the 22nd February 2011 Christchurch earthquake, is presented and discussed.

4.1 Canterbury earthquake sequence intensity measures

The observed maximum horizontal PGA isoseismals for both the Darfield and Christchurch earthquakes have been downloaded from the New Zealand Geotechnical Database (NZGD) (EQC, 2016), based on work by Bradley and Hughes (2012). The NZGD is a web-based repository for sharing geotechnical data relevant to the assessment of earthquake risk in the region. It contains observations from the Canterbury earthquake sequence and is the source of most of the observed data in this research for the Darfield and Christchurch earthquakes. The original PGA isoseismals are published at 0.01 g intervals, but are shown at 0.10 g intervals in Figure 4.1 for clarity. Although there was an intention to include corresponding maximum horizontal PGV isoseismals in the NZGD (Bradley, personal communication), this work has never been completed and so PGV isoseismals have been obtained from the USGS ShakeMap archive (USGS, 2015a, 2015b). The original PGV isoseismals are at 2 cm/s intervals, but are shown in Figure 4.2 at 10 cm/s intervals for clarity. Tonkin and Taylor, geotechnical engineering consultants to the EQC (van Ballegooy et al., 2014) have provided a GIS dataset containing information about the location of over 70,000 investigative boreholes in the city, with attribute information describing the qualitative surface land damage category at each location for both earthquakes. The land damage attribute has been designated based on a combination of post-earthquake on-ground surveys and aerial photography.

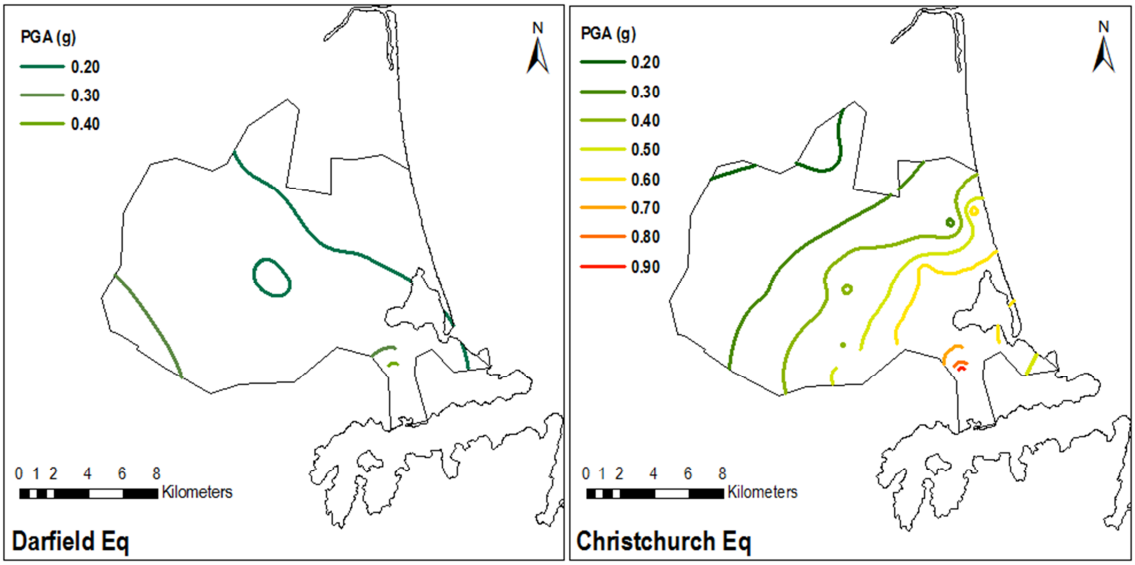

Figure 4.1 – Peak ground acceleration (PGA) maps for the Christchurch urban area from the Darfield and Christchurch earthquakes, based on data from the New Zealand Geotechnical Database

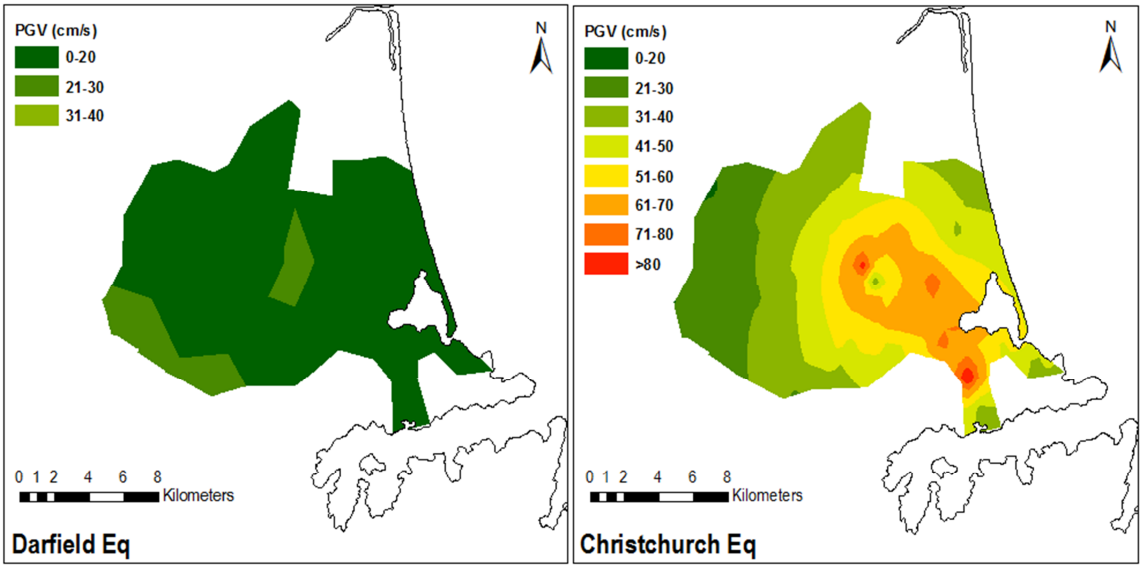

Figure 4.2 – Peak ground velocity (PGV) maps for the Christchurch urban area from the Darfield and Christchurch earthquakes, based on data from the US Geological Survey.

There are six land damage categories, which are listed and described in Table 4.1. Land damage category 2 is described by Tonkin & Taylor as ‘minor ground cracking’, reflecting the fact that no liquefaction ejecta material is observed on the surface. However, even when no ejecta material is observed, ground cracking can be interpreted as evidence of liquefaction in deeper soil layers. In subsequent studies of liquefaction in the Canterbury earthquake sequence, this category is described as either ‘liquefaction, certain’ (Brackley, 2012), which is defined as being

greatly affected by liquefaction, or 'marginal liquefaction' (Green et al. 2014; Maurer et al. 2014). Van Ballegooy and Malan (2013) go further and use observations of ground cracking as evidence of the direction, magnitude and extent of lateral spreading, implying that it is at the very least a precursor to lateral spreading. Almost all observation of category 2 are in very close proximity to other liquefaction observations and to waterways. So for the purposes of this research, category 2 is considered to describe lateral spreading.

Table 4.1 – Land damage categories in data provided by Tonkin & Taylor for qualitative liquefaction observations

Land damage category	Description
1	No liquefaction
2	Minor ground cracking
3	Liquefaction – moderate settlement only
4	Liquefaction – severe settlement only
5	Liquefaction – moderate lateral spreading
6	Liquefaction – major lateral spreading

Based on the data provided by Tonkin and Taylor, Figure 4.3 shows the approximate extents of liquefaction observations in both earthquakes. The extents have been extrapolated from the borehole sample points by applying Thiessen polygons (de Smith et al. 2009), which is a type of nearest neighbour analysis. In the Thiessen polygon method, discrete sampled point observations of a variable can be extrapolated to a surface of discrete zones by assigning locations in the unsampled space with the attributes of the closest sample point. For example, if the closest sample point to an unsampled location is observed to be l, then the unsampled location is assumed to be in land damage category 4 also. Maps of observed horizontal ($PGDf_H$) and vertical ($PGDf_V$) permanent ground deformations, measured using LiDAR technology, have been acquired from the New Zealand Geotechnical Database. LiDAR is a technique in which a laser scanner, fires rapid pulses of laser light towards a target object and then uses a light sensor to measure the distance between the scanner and the object based on the time taken for the pulse

to return, given that the speed of light is constant (LiDAR-UK, 2016). When this is repeated multiple times in quick succession, a complex 3D map of the surface of the target object can be constructed.

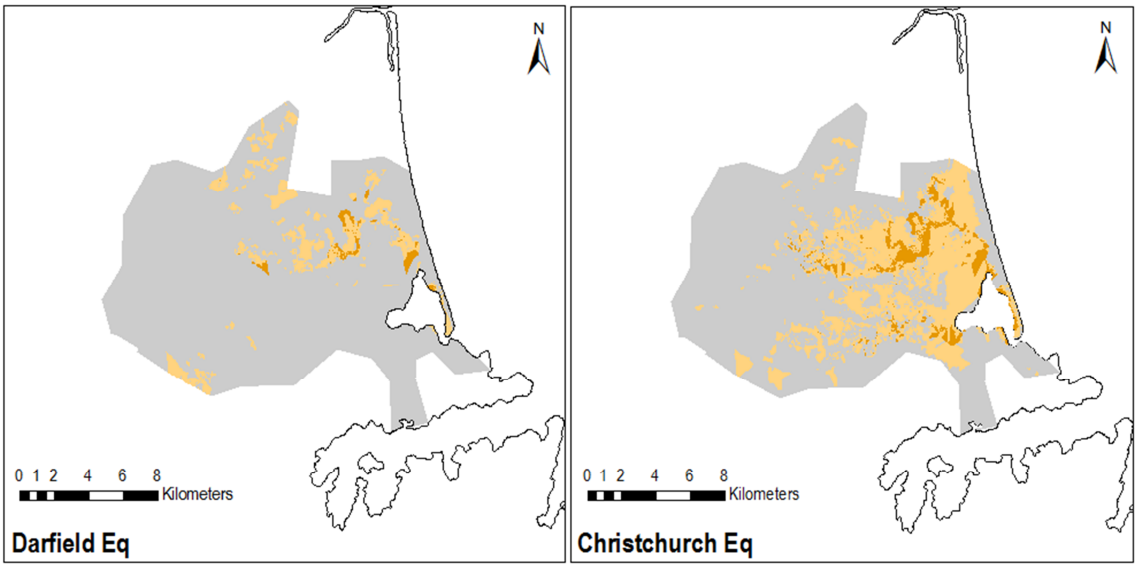

Figure 4.3 – Surface liquefaction observations in the Christchurch urban area due to the Darfield and Christchurch earthquakes based on sample data collected by Tonkin and Taylor. The maps indicate areas of no liquefaction (grey), vertical settlement (orange) and lateral spreading (brown)

In Christchurch, airborne LiDAR systems have been used to construct digital elevation models (DEMs) of the ground surface as rasters at a 5 m cell resolution. The first survey took place prior to the earthquake sequence in 2003 and has subsequently been repeated after the Darfield and Christchurch earthquakes. The differences between the DEMs from post-Darfield earthquake survey and the 2003 survey represent the vertical movement due to the Darfield earthquake, and similarly the differences between the DEMs from the post-Christchurch earthquake and the post-Darfield earthquake surveys represent the movement due to the Christchurch earthquake. In addition to liquefaction, elevation changes recorded by LiDAR include changes caused by tectonic uplift. Therefore, to evaluate the vertical movement due to liquefaction effects only, (i.e. the total settlement), the differences between LiDAR surveys have been corrected to remove the effect of the tectonic movement.

Figure 4.4[1] shows the total settlements after the Darfield earthquake. After the Christchurch earthquake, the condition of a cable is dependent on the cumulative effects of liquefaction from both earthquakes rather than just from Christchurch earthquake in isolation. Therefore Figure 4.5[1] shows the cumulative total settlements after the Christchurch earthquake. Horizontal movements have been estimated using a pattern-matching co-registration process (Leprince et al. 2007), also known as subpixel correlation, to find the relative position of corresponding pixels across successive DEMs (van Ballegooy et al. 2014). Figure 4.6[1] shows the horizontal movement after the Darfield earthquake and the cumulative horizontal movement after both the Darfield and Christchurch earthquakes.

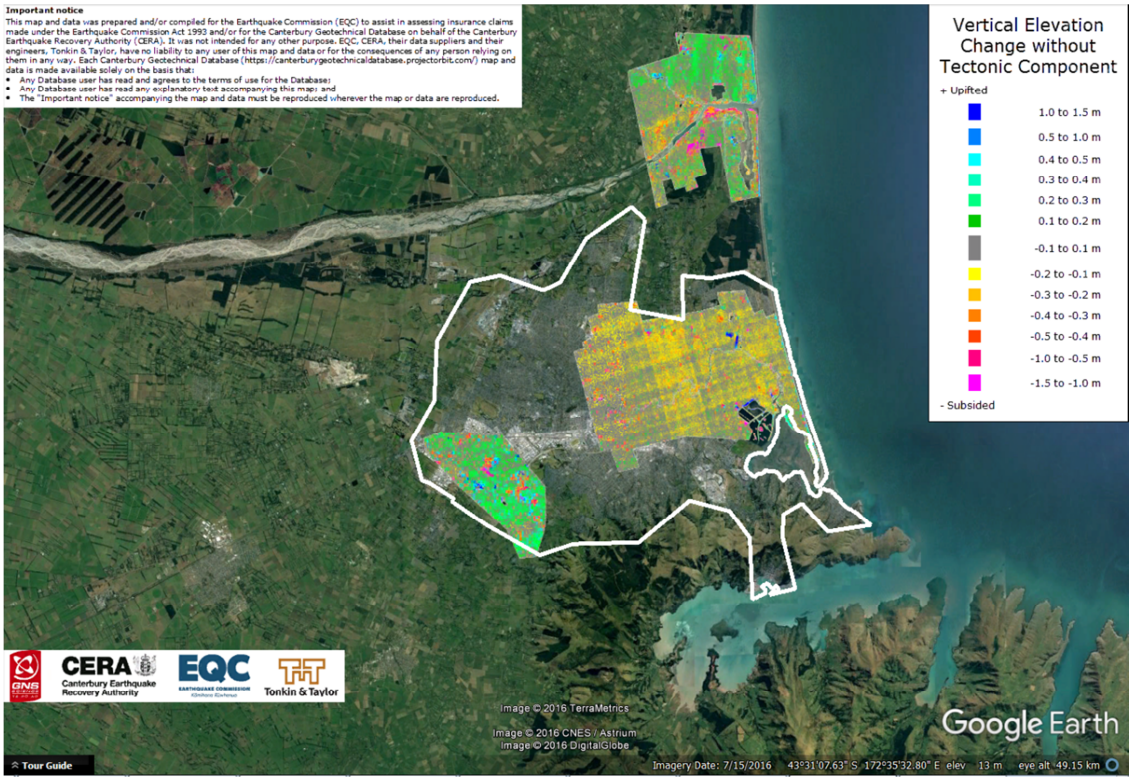

Figure 4.4 – LiDAR measurements of liquefaction-induced vertical settlement after the Darfield earthquake

[1] Figure 4.4, Figure 4.5 and Figure 4.6 were created from maps and/or data extracted from the Canterbury Geotechnical Database (https://canterburygeotechnicaldatabase.projectorbit.com), which were prepared and/or compiled for the Earthquake Commission (EQC) to assist in assessing insurance claims made under the Earthquake Commission Act 1993. The source maps and data were not intended for any other purpose. EQC and its engineers, Tonkin & Taylor, have no liability for any use of the maps and data or for the consequences of any person relying on them in any way. This "Important notice" must be reproduced wherever Figure 4.4, Figure 4.5 and Figure 4.6 or any derivatives are reproduced.

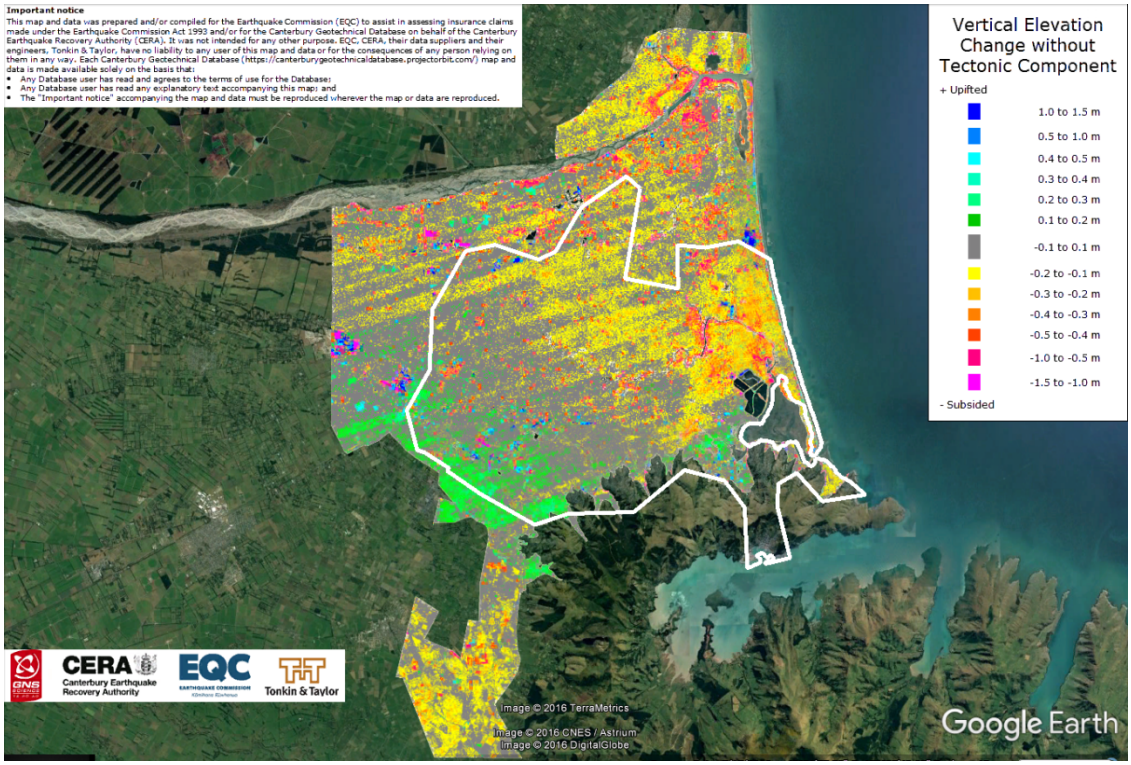

Figure 4.5 – LiDAR measurements of cumulative liquefaction-induced vertical settlement after the Darfield and Christchurch earthquakes

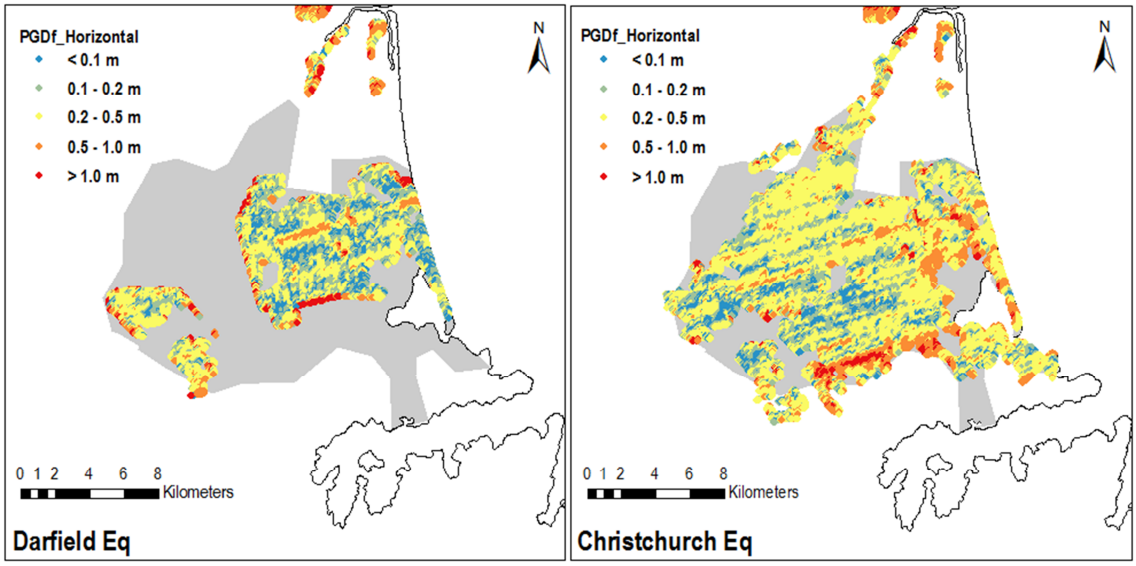

Figure 4.6 – Maps of horizontal ground movements ($PGDf_H$) after the Darfield earthquake and cumulatively after the Christchurch earthquake from LiDAR surveys. The maps have been reproduced from data from the New Zealand Geotechnical Database

It can be seen that the extents of liquefaction identified in Figure 4.3 do not correspond exactly to the extents of permanent ground deformations identified in Figure 4.4 to Figure 4.6. One reason is that the land damage attributes associated with borehole samples only represent observations of surface liquefaction and do not account for subsurface liquefaction, which could be measured by LiDAR. The exception is the assumption that land damage category 2 (minor ground cracking) is representative of subsurface lateral spreading. The use of Thiessen polygons to extrapolate from the sample observations adds further uncertainty. Also the LiDAR method for measuring ground deformations has some shortcomings. Metadata provided by the LiDAR contractor indicates accuracy of up to ±0.15 m in the vertical direction and up to ±0.55 m in the horizontal direction. To put this into context, the range of measured ground movements is up to ±1.5 m in the vertical direction and up to 3.2 m in horizontal direction. Furthermore, the pre-earthquake LiDAR survey took place seven years prior to the Darfield earthquake. Without intermediate surveys to identify and reconcile potential changes to elevation and position that may have occurred during the intervening period, it is assumed that all changes identified by the post-Darfield earthquake survey are due to liquefaction effects in that event, which may be an overestimation. Furthermore, the LiDAR analysis may not estimate the magnitude of permanent ground deformations with high precision. As a consequence it may yield false positive observations of liquefaction, i.e. measuring ground movements in locations where no liquefaction occurred. Therefore in this research, for analysis involving observed permanent ground deformations (e.g. developing new empirical fragility functions), it is proposed to validate the LiDAR measurements with the surface-level observations from the borehole data provided by Tonkin and Taylor.

To validate the LiDAR measurements, the study area is divided into four observed liquefaction ‘zones’, based on the land damage categories, as shown in Table 4.2. The four zones are: A) no liquefaction (category 1); B) observed liquefaction (categories 2 to 6); C) observed liquefaction with settlement only (categories 3 and 4); and D) observed liquefaction with lateral spreading (categories 2, 5 and 6). The zones are not exclusive since zones C and D are sub-divisions of

zone B. For empirical fragility analysis, an asset is only considered to have been exposed to lateral spreading if it is both located in Zone D and the LiDAR measurement shows a horizontal movement. An asset is only considered to have been exposed to settlement if it is both located in Zone C and the LiDAR measurement shows a vertical movement, All other assets are assumed to have not been exposed to liquefaction.

Table 4.2 – Composition of observed liquefaction zones in terms of land damage categories

Land damage category	Description	Zone A	Zone B	Zone C	Zone D
1	No liquefaction	Yes	No	No	No
2	Minor ground cracking	No	Yes	No	Yes
3	Liquefaction – moderate settlement only	No	Yes	Yes	No
4	Liquefaction – severe settlement only	No	Yes	Yes	No
5	Liquefaction – moderate lateral spreading	No	Yes	Yes	Yes
6	Liquefaction – major lateral spreading	No	Yes	Yes	Yes

4.2 Step 1 – System classification

This section sets out the classification of assets within the electric power network and water supply system in Christchurch. This is part of the exposure module of seismic risk assessment and relates to step 1 in the seismic risk assessment methodology presented in Figure 3.2. For each system, the description is divided into three parts: a general operational overview and detail of asset inventory for the system prior to the Canterbury earthquake sequence; a descriptive summary of damage suffered by the system in the Darfield and Christchurch earthquakes; and a description of changes to the system that have occurred since the Canterbury earthquake sequence. The final part represents the system classification for the seismic risk assessment model, but the first two parts are included as this information is needed for asset fragility function derivation/validation in section 4.5.

4.2.1 Electric power network

4.2.1.1 Pre-earthquake

This section describes the electric power network before the Canterbury earthquake sequence and, unless otherwise stated, is based on information from the most up-to-date Orion Asset Management Plan (Orion, 2009) at the time of the Darfield earthquake. Electric power supply in New Zealand is operated as a hierarchical network. Transpower is the national supplier, transmitting power along high voltage lines from generation sites to demand centres, where it is further transmitted and distributed by local suppliers to customers. Power generation for Christchurch is primarily from hydro-electric schemes in the Alpine Lakes region, 250 km to the south-west of the city and a seismically active location. However, New Zealand operates a National Grid, meaning that in the event of the Alpine plants suffering outages, Christchurch can still receive power from generation plants in other parts of the country including geothermal sources on North Island. This redundancy is important not only because of the risk of earthquakes, which are relatively rare, but also to mitigate the impact of the more common dry periods, which prevent hydro-electric plants from generating at capacity. Because of this redundancy and the extremely low probability that all generation sources will fail simultaneously due to the same event, generation plants are not included in the analysis of the Christchurch network, which focuses on substations and conduits.

Orion is the local company responsible for sub-transmission and distribution of power to customers in Christchurch and the nearby suburbs and rural settlements. They receive power from Transpower at five substations known as Grid Exit Points (GXPs), where the power is transformed from 220 kV down to medium voltages (11 kV or 33 kV or 66 kV) and then re-distributed into the sub-transmission network. At the next level in the network hierarchy, Orion operates 50 district/zone substations (30 in the Christchurch urban area) that convert power to 11 kV (if required) and re-distribute into the 11 kV sub-transmission network. There are then a further 271 network substations that do not transform voltage but simply re-distribute supply into the 11 kV distribution network. It should be noted that in some parts of Christchurch, the

distribution network is bypassed and power is carried directly from district/zone substations to the final level of the hierarchy, the low voltage network, where Orion own over 10,000 distribution substations that convert power to 400 V and then supply directly to customers. A schematic explaining the interactions between levels in the network hierarchy is shown in Figure 4.7.

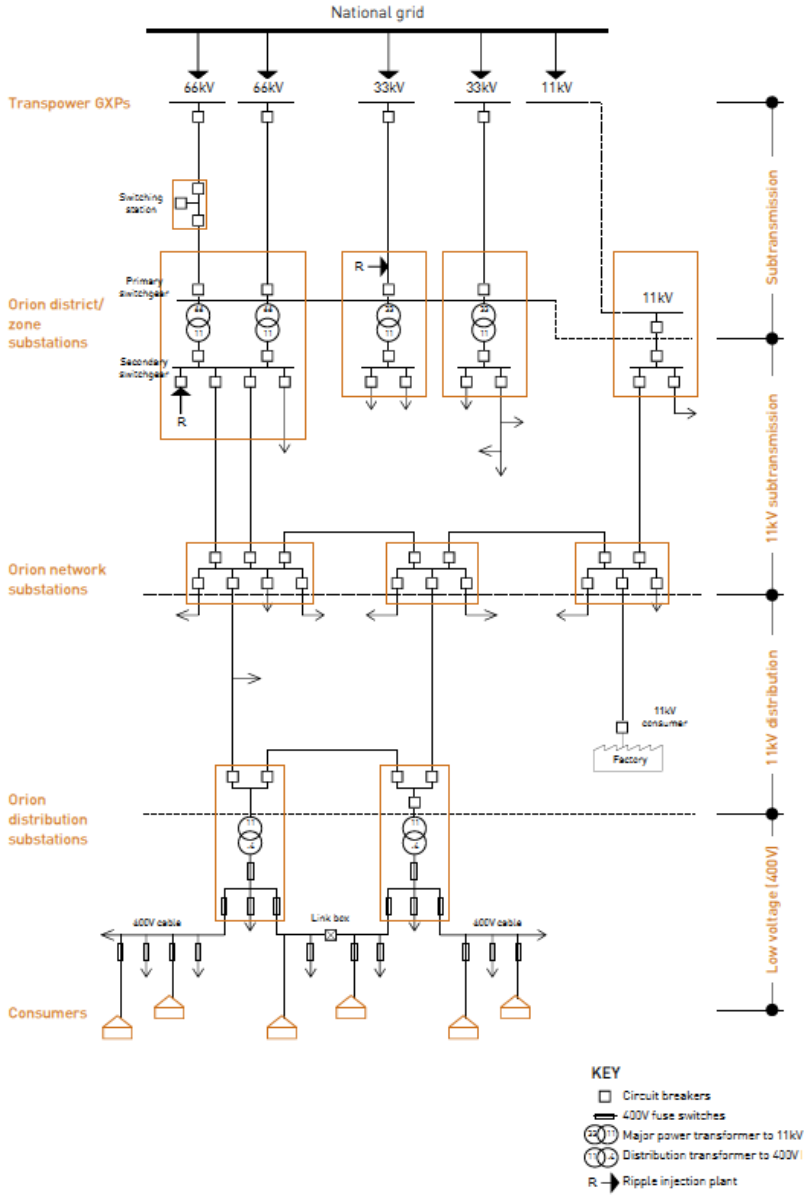

Figure 4.7 – Schematic of Orion network hierarchy in Christchurch urban area (source: Orion Asset Management Plan, 2009)

Around 70% of the conduits within Christchurch are buried cables, equating to 4,718 km. Buried cables can be categorised by their insulation type and their conduction material, although since the insulation provides the structure to a cable, it is thought that insulation is more susceptible to damage from ground movements (Orion, personal communication). The predominant insulation types in Christchurch are: oil-filled cables; paper-insulated lead-covered armoured cables (PILCA); PILCA cables reinforced with high density polyethylene (PILCA HDPE); cross-linked polyethylene cables (XLPE); and polyvinyl chloride cables (PVC). Conduction materials used in Christchurch are copper and aluminium. There are 2,176 km of overhead lines in the city but 97% are low voltage 400 V and the remaining 59 km are 66 kV lines in the sub-transmission network.

Orion has only provided the geospatial data required for analysis for the sub-transmission network and so the analysis only considers assets upstream of (and including) the district/zone substations, as shown in Figure 4.8, excluding the Moffett and Shands substations, which do not serve customers in the urban area of Christchurch. At lower levels in the hierarchy of an infrastructure network, there is greater redundancy in supply paths and hence a lower level of functional vulnerability, i.e. the network at this level can accommodate a larger number of faults without affecting functionality (Vanzi, 1996; Hwang et al., 1998). Therefore, whilst excluding the lower voltage parts of the network will have some effect on system performance measurements, it is likely to be a minor one. The distribution of types of conduits in the analysis zone is summarised in Table 4.3.

Table 4.3 – Summary of electrical conduits managed by Orion in Christchurch urban area in 2009 from GIS data

Voltage	Type	Total Length	Insulation Material
66 kV	Overhead line	59 km	n/a
66 kV	Buried cable	64 km	58 km oil-filled 6 km XLPE
33 kV	Overhead line	15 km	n/a
33 kV	Buried cable	27 km	18 km XLPE 5 km PILCA 4 km oil-filled
11 kV	Buried cable	75 km	56 km PILCA 15 km XLPE 1 km PILCA HDPE 3 km unspecified

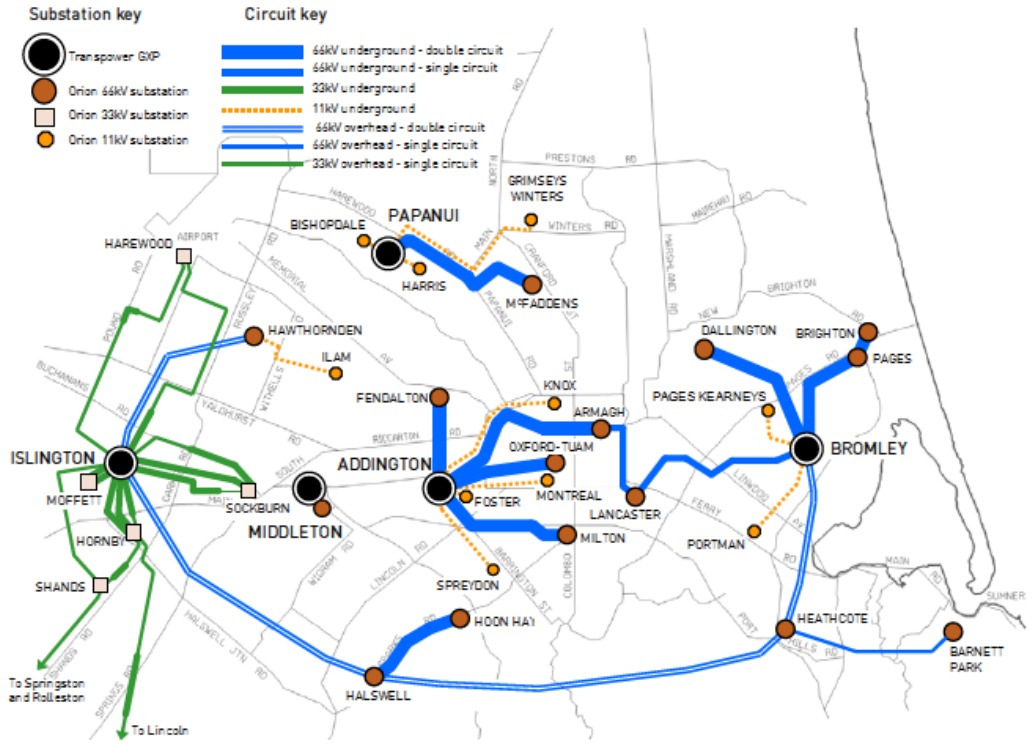

Figure 4.8 – Orion sub-transmission network prior to the Canterbury earthquake sequence (source: Orion Asset Management Plan, 2009)

4.2.1.2 Electric power network performance

Electric power infrastructure generally performed well in the Darfield earthquake. Minor damage to transmission GXPs included: damage to a line at Papanui; transformer banks tripping at Islington; and a leaning transmission tower at Bromley. Two of Orion's district/zone substations at Pages Road and Brighton experienced minor physical damage but remained operational and the impact on Orion's network was caused primarily by transformers tripping (Kwasinski et al., 2014). Orion carried out 24 repairs on 11 kV buried cables but there were no direct faults on buried 66 kV cables or on overhead lines. However, two 66 kV cables were crushed at the Dallington Bridge, as the bridge failed due to liquefaction at the transition point between the bridge and the adjoining ground. Almost 170,000 of Orion's customers (from a total of 193,000) lost power immediately after the Darfield earthquake, although by the end of the day, just over 15,000 customers remained without power. In total, Orion estimate 90 million customer minutes of outages after the Darfield earthquake (Kwasinski et al. 2014).

In the Christchurch earthquake, Transpower infrastructure was not greatly affected with damage concentrated at the Bromley GXP, where damage occurred to switchgear in the control building and a voltage transformer in the switchyard (Kwasinski et al., 2014). The level of damage at Bromley was low considering the severity of liquefaction in the area. None of the district/zone substations were damaged. Only one of the network substations, at St. Andrew's Hill, suffered damage due to ground shaking while one other suffered damage due to rock fall in Sumner, at the foot of the Port Hills (Eidinger and Tang 2012). 50% of 66 kV cables experienced faults, compared to 15% of 11 kV and just 0.6% of 400 V cables (Kwasinski et al, 2014). Immediately after the Christchurch earthquake, over 70% of the city did not have access to power, in part because automatic shutdown systems functioned as designed, preventing short circuits and potential earthquake-induced fires (Giovinazzi et al., 2011). Most of the outages were caused by liquefaction damage to cables, while above-ground components, including substations and overhead lines performed well (Kwasinski et al., 2014). 50% of customers had power restored the same day; 75% of customers had power restored within four days and after two weeks, 98%

of customers were receiving power (Giovinazzi et al. 2011). In total, Orion estimate 629 million customer minutes of outages after the Christchurch earthquake (Kwasinski et al. 2014). The map in Figure 4.9 shows the approximate extent of the area affected by power outages.

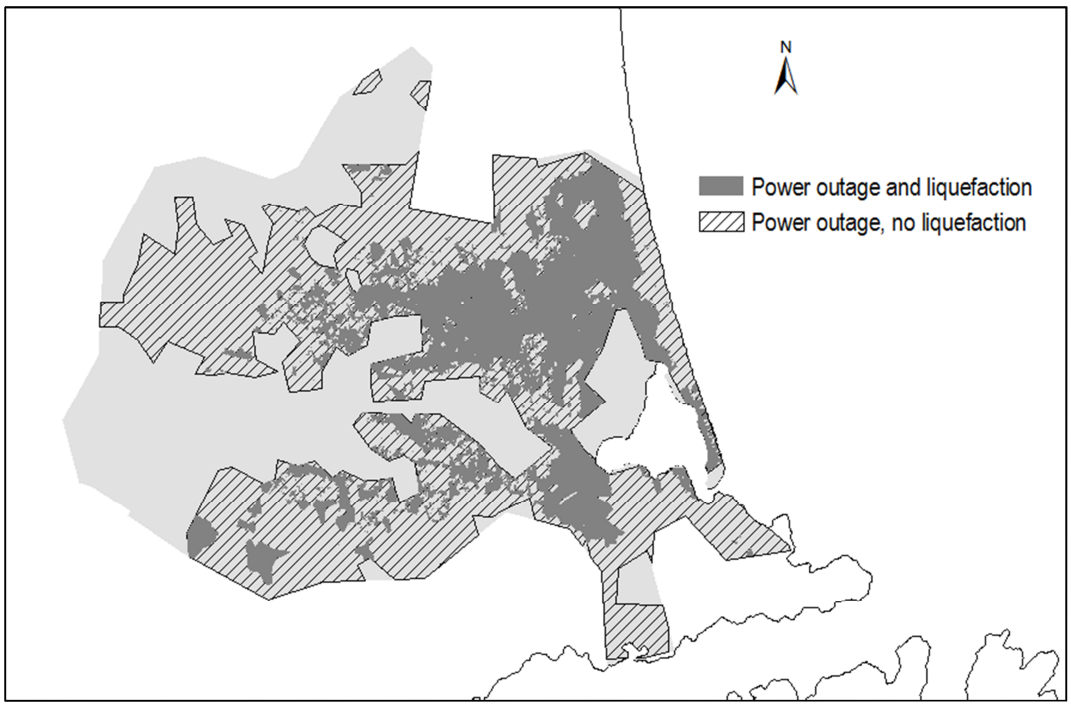

Figure 4.9 – Map showing approximately the areas affected by power outages immediately after the Christchurch earthquake

The good performance of the substations in both earthquakes can be traced back to the decision by Orion in the 1990s to seismically reinforce all substations at a cost of 6 million NZD. This followed lessons learned from the M_W 6.5 earthquake that struck Edgecumbe on North Island in 1987 (Giovinazzi et al., 2011). Buried cable damage was found to be the most costly type of damage to the power system and the main reason for long outages after the February 2011 earthquake (Kwasinski et al., 2014). Although 66 kV cables were significantly affected by the earthquakes, Orion has provided geospatial damage data only for the 11 kV cables. The location of 11 kV cable repairs in relation to areas where liquefaction was observed is shown in Figure 4.10. It can be seen that most of the cable damage occurred in areas of liquefaction and typical examples of the type of damage caused by liquefaction are shown in Figure 4.11.

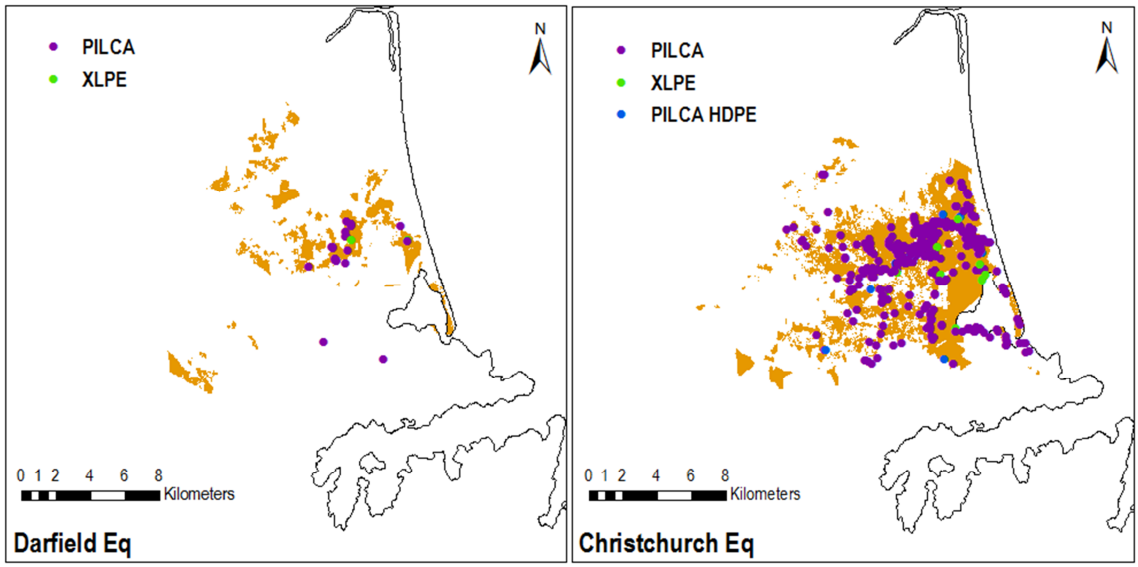

Figure 4.10 – Locations of recorded buried 11kV cable faults due to each earthquake by insulation material, mapped against areas of observed liquefaction (brown)

Figure 4.11 – Examples of typical curvature damage observed amongst buried cables due to the Canterbury earthquakes (photos courtesy of Andrew Massie at the Christchurch Polytechnic Institute of Technology)

The average fault rate for 66 kV cables over the two earthquakes was 0.11 faults per kilometre (Kwasinski et al., 2014) – since 1994 there have been no faults amongst these cables due to plant failure or third party effects (Orion 2009). The average fault rate for 11 kV cables over the two earthquakes is 0.12 faults per kilometre (calculated independently by the author from geospatial data provided by Orion). This compares to pre-earthquake observed annual repair rates of less than 0.03 faults per kilometre due to plant failure or third party effects (Orion 2009). Kwasinski et al. (2014) observe that both copper and aluminium should be able to accommodate the moderate extension that could be expected due to liquefaction and therefore

this significant cable faulting is more likely to be caused by yielding of the outer insulation layer. Analysis of the geospatial data indicates that across both earthquakes, PILCA cables exhibited a repair rate of 0.143 faults per km, whilst XLPE cables exhibited a repair rate of 0.022 faults per km. More detailed investigative analysis of cable fault rates is described in section 4.5.2.

4.2.1.3 Post-earthquake

Since the Canterbury earthquake sequence, a small number of changes have been made to the Christchurch power network. Some of these changes have been due to earthquake damage but some have also been due to general network rationalisation and performance improvement. This section briefly describes these changes based on information the latest Orion Asset Management Plan (Orion, 2014). The general hierarchy of the system is the same although there is a minor change in nomenclature. The level previously identified as the 11 kV sub-transmission network is now known as the 11 kV primary distribution network; and the level previously known as the 11 kV distribution network, is now known as the 11 kV secondary distribution network. Orion has only provided the geospatial data required for analysis for the sub-transmission network and so the analysis only considers assets upstream of (and including) the district/zone substations, as shown in Figure 4.12. A major change is that the Papanui GXP was taken over by Orion and now operates as an 11 kV district/zone substation, which receives power from the Islington GXP. This has led to the addition of three 66 kV lines to the Orion network, linking the two substations, and to the closure of the nearby Harris district/zone substation. The Brighton and Pages Road district/zone substations permanently closed after the Christchurch earthquake due to liquefaction damage and have been replaced by a new substation at Rawhiti. Much of the underground cabling that led out from the Bromley GXP was also damaged by liquefaction and so the Dallington and new Rawhiti substations currently receive power from Bromley GXP across temporary overhead lines. A new district/zone substation has also been constructed at Prebbleton but this does not serve the urban area of Christchurch and so is not included in the

analysis. The current inventory for conduits in the sub-transmission network is summarised in Table 4.4.

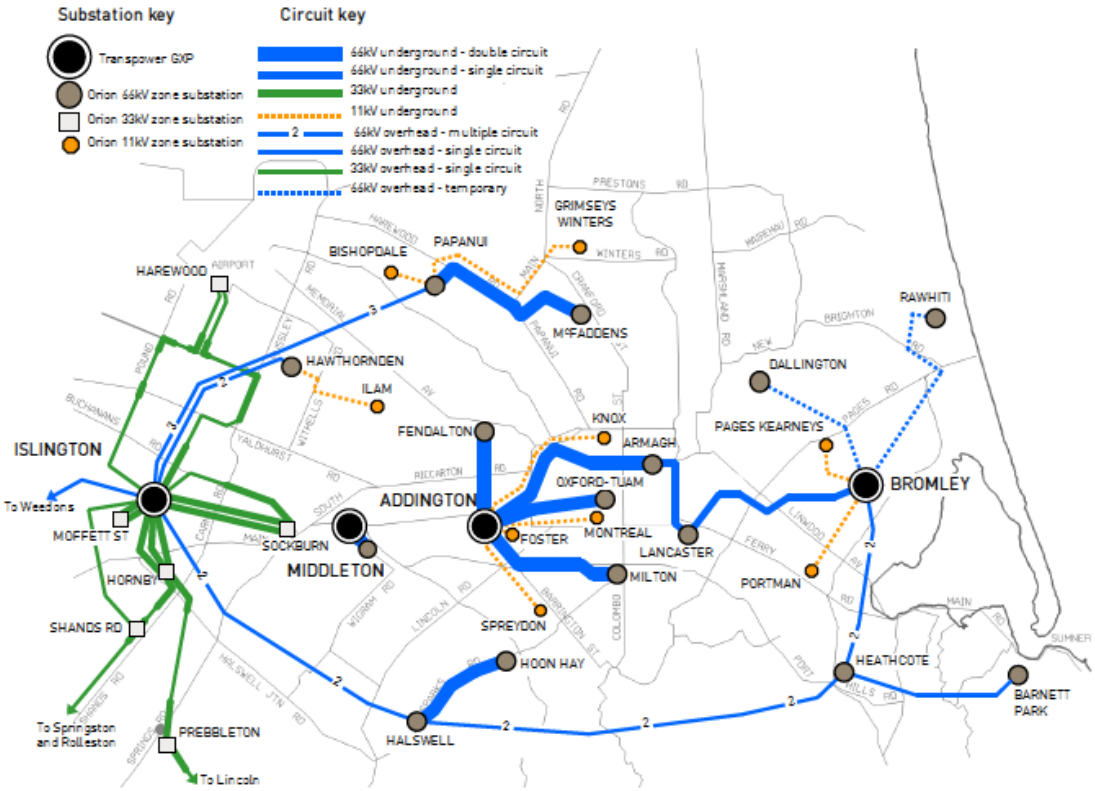

Figure 4.12 – Orion sub-transmission network following changes implemented after the Canterbury earthquake sequence (source: Orion Asset Management Plan, 2014)

Table 4.4 – Summary of electrical conduits managed by Orion in Christchurch urban area in 2014 from GIS data

Voltage	Type	Total Length	Insulation Material
66kV	Overhead line	90km	n/a
66kV	Buried cable	56km	41km oil-filled 15km XLPE
33kV	Overhead line	7km	n/a
33kV	Buried cable	14km	all XLPE
11kV	Buried cable	75km	56km PILCA 15km XLPE 1km PILCA HDPE 3km unspecified

4.2.2 Water supply system

4.2.2.1 Pre-earthquake

Christchurch City Council (CCC) is responsible for the water supply in the Christchurch urban area and surrounding fringe settlements. Water is obtained by pumping through wells from underground gravel aquifers, which extend under the city and westwards under the Canterbury Plains, to depths of up to 220 m. Prior to the Canterbury earthquake sequence, approximately 50 million cubic metres of water was abstracted per year to supply customers (CCC, 2005). The abstracted water is of high quality and requires no additional treatment. The Christchurch water supply system is divided into ten pressure supply zones (seven in the urban area). Each zone contains a number of wells and associated primary pumping stations to abstract water from aquifers and a pipe network to carry water to customers. The majority of pipes in the city operate by gravity in some of the supply zones also contain additional booster pumping stations to carry water to properties on higher ground. The zones operate independently, meaning that a customer can only receive water from a well in the same zone. A map of the CCC water supply zones is shown in Figure 4.13.

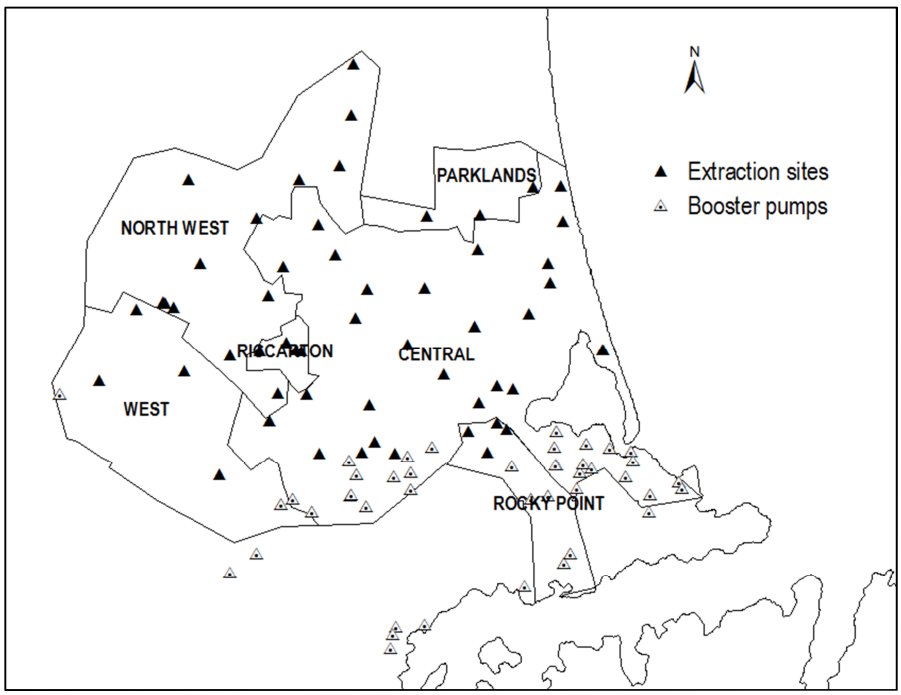

Figure 4.13 – Map of water supply pressure zones and extraction sites in Christchurch

The University of Canterbury acts as a custodian and access manager for CCC's GIS data service. The geospatial water supply system datasets used in this study have been acquired through this resource rather than from CCC directly and the following summary of network assets is based on these geospatial datasets. There are 56 extraction sites in the Christchurch urban area, housing a total 105 primary pumping stations, of which 43 had access to back-up power supply in the form of a diesel generator. Some pumping stations are linked to more than one well and there are a total of 147 wells. The system is supported by 59 booster pumps, spread across 41 sites, to provide water to the Port Hills area to the south of the city. This includes nine sites that are outside of the urban area but which receive water from one of the six urban pressure zones. None of the booster pumping stations had access to back-up power. The locations of the extraction sites and booster pumps are shown in Figure 4.13, with reference to their supply zones.

There is a total of 3,357 km of pipes in the Christchurch water system of which 53 km is classified as 'trunk mains', 1,686 km is classified as 'mains' and 1,618 km classified as 'submains'. Apart from size – 93% of trunk mains have a diameter of at least 300 mm, whilst 93% of mains have a diameter of less than 300 mm – the distinction between 'trunk main' and 'main' is not clear. The extent of the trunk main network only covers a small part of the city and not every supply zone contains a trunk main. Furthermore, the majority of primary pumping stations pump directly into a mains pipe. Although the question has been put to CCC, no response has been forthcoming. Therefore in this study trunk mains and mains are assumed to be topologically equivalent so that there are two levels in the water network hierarchy: trunk mains/mains and submains. The total network consists of almost 120,000 individual pipes and modelling both the fragility and connectivity of all of these will result in a very long computation time even for just a single event and so may be prohibitive for a Monte Carlo simulation. It is therefore proposed to use a simplified network. Even restricting the analysis to only trunk mains and mains leaves almost 50,000 pipes, so a stricter condition is needed. The network is initially thinned by only including pipes with diameter >300 mm. However, this

results in a disjointed network, meaning that there are sections of pipe or booster pumps that are not connected to a well, because water is carried through smaller diameter pipes to reach them. It is therefore proposed to include some of the smaller diameter pipes in the simplified network, to ensure that all large diameter pipes (>300 mm) and booster pumps are connected to a well. In some cases there is more than one route option to link an isolated asset to the rest of the network. When this arises, the route with the largest diameter is pipe selected if possible, or if there is no difference in diameter then the shortest route is selected. The final analysis network consists of 6,161 pipes. A range of materials are used to construct pipes in Christchurch and the quantities of each material in the analysis network are shown in Table 4.5.

Table 4.5 – Lengths of pipe in pre-earthquake analysis network by construction material

Material	**Trunk**
Asbestos Cement (AC)	108
Cast Iron (CI)	23
Cement Lined Ductile Iron (CLDI)	3
Cement Lined Steel (CLS)	36
Ductile Iron (DI)	18
Galvanised Steel (GALV)	<1
High Density Polyethylene (HDPE)	<1
Medium Density Polyethylene (MDPE80)	<1
Modified Polyvinyl Chloride (MPVC)	14
Polyvinyl Chloride (PVC)	31
Steel (STEEL)	16
Unplasticised Polyvinyl Chloride (UPVC)	15
Other	2
Unknown	1
Total	267

4.2.2.2 Water supply system performance

In the Darfield earthquake a booster pumping station building was damaged due to ground shaking. Eight wells failed and a further one was damaged, all due to liquefaction. Some wells and booster pumping stations that did not have back-up generators lost power immediately after the earthquake, but since there was sufficient water available to meet local demand, this did not affect service provision (Eidinger and Tang 2012). Pipes in liquefaction areas were severely damaged with approximately 280 repairs required, mostly completed within a week. The failure of pipes and wells resulted in some loss of pressure in the water system, but fortunately fire was not a major factor after the earthquake so this had no significant consequences. In the Christchurch earthquake, 20 wells were damaged, in some cases inundating local streets with water (Giovinazzi et al. 2011).The February earthquake resulted in almost 4,000 pipe repairs spread across the city and took six weeks to complete (Eidinger and Tang, 2012). It was notable that at Horseshoe Lake and Lyttelton, old asbestos cement and cast iron pipes that were damaged in the Darfield earthquake were replaced immediately by new HDPE pipes and subsequently no damage was reported in the new HDPE pipes following the Christchurch earthquake.

The geospatial data acquired from the University of Canterbury includes information on pipe damage observed in the Christchurch earthquake. More detailed investigative analysis of this data is described in section 4.5.4 but a summary of the key points is presented here. The data indicate that about one third of mains and submains repairs occurred in AC pipes, which is not surprising given the high exposure of this material. However, a similar number of repairs occurred on GALV pipes despite a much lower exposure and this exhibited the highest repair rates of all materials. Amongst the materials that are more prevalent in the mains network, AC pipes exhibited the highest rate in liquefaction areas, whilst CI pipes exhibited the highest rate in non-liquefaction areas. In both areas, AC and CI pipes had significantly higher repair rates than PVC or MPVC pipes. Overall, the repair rate for all pipes in liquefaction areas was 2.32 per km compared to 0.58 per km in non-liquefaction areas.

In the immediate aftermath of the Christchurch earthquake, 80% of Christchurch lost connection with mains water supply. For the first few days afterwards, this figure reduced to about 50% and more than a third of households were without water for over a week (Ministry of Health, 2012). To increase the supply of fresh water, two temporary desalination plants were set up by the army at Lyttelton and New Brighton and tankers used to distribute water. Within a month over 95% of occupied units had water, but a 'boil order' was in place for over six weeks for most of the city due to potential contamination caused by severe damage to the wastewater system. As a precaution, portable chlorination stations were set up across the city. The boil orders were only removed once the city experienced 14 consecutive days without water quality issues. Although pumping stations generally performed well owing to pre-earthquake seismic engineering upgrades, water conservation orders were put into place as a precautionary measure (Billings and Charman 2011).

4.2.2.3 Post-earthquake

Since the Canterbury earthquake sequence, a number of changes have been made to the Christchurch water supply network. Changes to the wells and primary pumping stations have been found by searching the Environment Canterbury online consent database (ECan, 2014a), for current water extraction permits in the Christchurch urban area. Six of the sites from which water was previously extracted are no longer in use. These are: Withells Road, Mandeville Street, Palmers Road, Chapman Road and two separate sites around the Lyttelton Tunnel Road. However, there are seven new extraction sites at: Hills Road, Mairehau Road, Raymond Road, Rawhiti Domain, Wilmers Road, Shands Road and Annex Road. In total there are currently 57 extraction sites housing 113 primary pumping stations and 153 wells. 46 of the pumping stations have back-up power. There have been no changes to the number or location of booster pumps. Changes to pipes are based on differences between the pre- and post-earthquake sequence geospatial datasets provided by the University of Canterbury. The total length of pipes in Christchurch has reduced after the earthquakes. This is due to the 'red zoning' policy of the Canterbury Earthquake Recovery Authority (CERA), which deemed certain suburbs in the east

of the city that had been severely affected by liquefaction to be uninhabitable and therefore abandoned. The current quantities of pipe by material in the analysis network are shown in Table 4.6.

Table 4.6 – Lengths of pipe in post-earthquake analysis network by construction material

Material	**Length (km)**
Asbestos Cement (AC)	109
Cast Iron (CI)	23
Cement Lined Ductile Iron (CLDI)	3
Cement Lined Steel (CLS)	36
Ductile Iron (DI)	18
Galvanised Steel (GALV)	<1
High Density Polyethylene (HDPE)	<1
Medium Density Polyethylene (MDPE80)	<1
Modified Polyvinyl Chloride (MPVC)	14
Polyvinyl Chloride (PVC)	31
Steel (STEEL)	16
Unplasticised Polyvinyl Chloride (UPVC)	16
Other	2
Unknown	1
Total	270

4.3 Step 2 – Seismic hazard model

4.3.1 Earthquake sources

As discussed in section 2.3, Monte Carlo simulation of scenario events is the most appropriate way of modelling risk to spatially distributed infrastructure systems. For the case study, the method of the New Zealand National Seismic Hazard Model (NSHM) (Stirling et al., 2012) is used as the basis for generating the Monte Carlo simulation events. This includes both active

fault and distributed seismicity sources within a 200 km radius of Christchurch. Geospatial data for active faults has been provided by GNS Science, the national geoscience research establishment for New Zealand, and has been supplemented by attribute data from Stirling et al. (2008) and for the Greendale Fault from Beaven et al. (2010). A total of 80 active faults exist within 200km of Christchurch, all of which are examples of shallow crustal seismicity (Stirling et al., 2008). The known attributes of theses faults include: location of surface projection, characteristic magnitude, fault style, length, strike, dip, width, depth to top of rupture and approximate return period for characteristic magnitude. For distributed seismicity, Stirling et al. (2012) identify 16 source zones for New Zealand, of which six contribute to the seismic risk of Christchurch. The attributes for each source zone are the fault type, the maximum magnitude (M_{cutoff}) and b_{GR} (the slope of the Gutenberg-Richter relationship). The locations of the active faults and extents of the source zones relevant to Christchurch are shown in Figure 4.14. The author has assigned names to the six source zones from the Stirling et al. (2012) study: North, West, North East, Mid-Canterbury, Christchurch and South. In addition, the author proposes a new distributed seismicity zone, Offshore East, which has not been considered by Stirling et al. (2012) but could generate large events that may affect Christchurch.

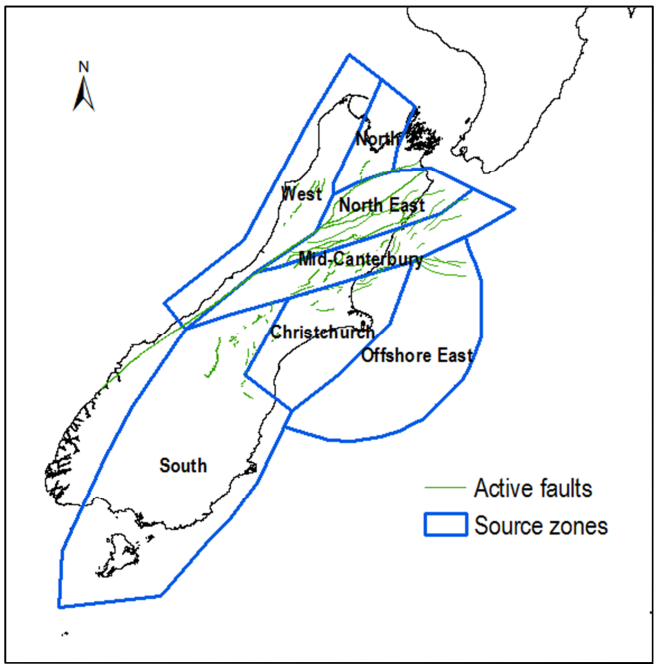

Figure 4.14 - Map of active faults and distributed seismicity source zones within 200km of Christchurch

The Gutenberg-Richter b-value (b_{GR}) for each zone has been estimated empirically by Stirling et al. (2012) from the observed earthquake catalogue up to 2009. The catalogue is complete from 1964 for magnitude, $M \geq 4$; from 1940 for $M \geq 5$; and from 1840 for $M \geq 6.5$. However, this means that events from the Canterbury earthquake sequence are not included in the calculation. Hence b_{GR} for each zone are re-calculated here using the approach stated by Stirling et al. (2012), but with a catalogue up to the end of 2014. For each zone, the Gutenberg-Richter relationship is determined on earthquakes with $M_W \geq 4$. However the earthquake catalogue dataset from GeoNet (GNS Science, 2014) includes earthquakes attributed with local magnitudes, M_L, or surface wave magnitudes, M_S. To convert M_S to M_W, the global empirical functions derived by Scordilis (2006) in Eq.'s 4.1 and 4.2 are employed.

$$\text{for } M_W \leq 6.1: \qquad M_W = 0.67M_S + 2.07 \tag{4.1}$$

$$\text{for } M_W > 6.1: \qquad M_W = 0.99M_S + 0.08 \tag{4.2}$$

To convert M_L to M_W, the New Zealand-specific empirical functions derived by Ristau (2009) in Eq.'s 4.3 and 4.4 are employed.

$$\text{for depth} \leq 33\text{km}: \qquad M_W = \frac{M_L - 0.73}{0.88} \tag{4.3}$$

$$\text{for depth} > 33\text{km}: \qquad M_W = \frac{M_L - 0.05}{1.09} \tag{4.4}$$

The next step is to remove aftershocks from the catalogue, which is done using the declustering algorithm of Reasenberg (1985). The algorithm identifies a spatial and temporal interaction zone around the first earthquake in the catalogue. The spatial extent of the earthquake interaction zone is based on an estimate of the stress redistribution around it. First the moment of the earthquake, M_0, measured in Nm, is calculated from the relationship proposed by Bakun (1984) in Eq. 4.5. Then the source dimension, a_{dim}, measured in km and defined as the radius of a circular crack corresponding to M_0 with a stress drop of 30 bars (Kanamori and Anderson,

1975), is calculated by Eq. 4.6. The spatial extent of the interaction zone is assumed to be 10 times the source dimension.

$$M_0 = 10^{(17+1.2M)} \times 10^{-7} \tag{4.5}$$

$$a_{\dim} = \sqrt[3]{\frac{7/16 M_0}{30 \times 10^5}} \times 10^{-3} \tag{4.6}$$

The temporal component, τ_{IZ}, of the earthquake interaction zone is estimated as the time period within which there is 0.95 probability of observing an aftershock, assuming that the aftershock sequence follows a time-dependent Poisson process. τ_{IZ} is calculated using Eq. 4.7, where t is the time interval in days between the earthquake around which the interaction zone is being determined and the previous earthquake in this equation (assumed to be 1 for the first event), M_{min} is the magnitude for which the earthquake catalogue is complete, which in this case is 4 and M_{max} is the maximum magnitude observed so far in the aftershock sequence, which for the first event is simply its own magnitude.

$$\tau_{IZ} = \frac{3t}{10^{2/3(M_{\max} - M_{\min} - 1)}} \tag{4.7}$$

The next earthquake is classified as an aftershock of the first earthquake if it falls within the spatial extent and time window of its interaction zone. For subsequent earthquakes, the location and time are tested against the interaction zone of the previous earthquake in the same way. However, if the previous earthquake has been classified as an aftershock, then the current earthquake being tested should also be tested against the interaction zone of the highest magnitude earthquake in that aftershock sequence. In this case the spatial extent of the interaction zone is assumed to be equal to the source dimension of the largest aftershock.

After declustering, the next step is to rectify issues with the depths attributed to some of the events. This applies to older events for which the location was poorly constrained and so a depth of either 5 km, 12 km or 33 km was arbitrarily assigned in order to find a hypocentre solution.

Rather than ignore these events, Stirling et al. (2012) use the McGinty (2001) method to statistically re-assign their depths. In general, earthquakes assigned a depth of 5 km are known to have occurred in the crust (top 10 km). Therefore depths for these events are re-assigned by randomly sampling between 0 km and 10 km assuming a uniform distribution. For events with an assigned depth of 12 km or 33 km, there is a two-stage process to re-assignment. The first stage determines whether the event is classified as shallow (≤35 km) or deep (>35 km). For each zone, earthquakes in the catalogue with known depths are used to empirically determine the proportion of shallow earthquakes in the zone. Then for each event with an assigned depth, a random number is generated by uniformly sampling between 0 and 1. The event is classified as shallow if the random number is less than or equal to the proportion of known shallow events, or classified as deep otherwise. The second stage determines the actual depth. For shallow events, depths are re-assigned by randomly sampling between 0 km and 35 km with a uniform distribution and for deep events, depths are re-assigned by randomly sampling between 35 km and 100 km with a uniform distribution. The final processed catalogue is used to determine new Gutenberg-Richter b-values for each zone by carrying out linear regression of the logarithm of the observed annual exceedance rate against magnitude at intervals of 0.1. The revised attributes of each source zone are shown in Table 4.7.

Table 4.7 – Revised distributed seismicity source zone attributes for catalogue up to 2014

Zone	$\mathbf{M_{cutoff}}$	**Fault type**	$\boldsymbol{b_{GR}}$
North	7	Reverse	1.08
West	7.7	Reverse	1.06
North East	7	Strike-slip	1.01
Mid-Canterbury	7	Reverse/Strike-slip	0.86
Christchurch	7	Reverse	1.14
Offshore East	7	Reverse	0.85
South	7	Reverse	0.93

Whilst values for b_{GR} to the entire zone, the corresponding Gutenberg-Richter a-value (a_{GR}) varies within each zone. For a-values, Stirling et al. (2012) model New Zealand as a three-dimensional grid of 0.1° x 0.1° cells (approximately 11 km in N-S direction and 8 km in E-W direction) with five depth layers: 10 km, 30 km, 50 km, 70 km and 90 km. Each cell and depth layer combination is treated as a distinct potential earthquake point source with its own Gutenberg-Richter coefficients (a_{GR}, b_{GR}), fault type and maximum magnitude corresponding to the source zone within which it is located. Two changes are made for this study. First, the grid is modelled with a 10 km x 10 km interval for simplicity and secondly, since Stirling et al. (2008) identified that all sources relevant to Canterbury are shallow, only two depth layers are used: 10 km and 30 km. With a 200 km buffer around Christchurch, there are 1,405 cells per layer and so 2,810 cells in total.

The procedure for calculating a_{GR} depends firstly on calculating the Gutenberg-Richter annual frequencies (N_{GR}) for each cell and smoothing the initial cell attributes (N_{GR}, b_{GR}, maximum magnitude) to prevent a situation where the value of an attribute changes abruptly between two adjacent cells (e.g. at the boundary between source zones). Using the same catalogue used previously to revise b_{GR} for each source zone, each observed event is assigned to a cell based on its location. For each cell, three values of N_{GR} are then determined: $N_{GR1} = N_{GR}(M \geq 4$ for 1964 to 2014); $N_{GR2} = N_{GR}(M \geq 5$ for 1940 to 1963); and $N_{GR3} = N_{GR}(M \geq 6.5$ for 1840 to 1939). The size of the catalogue mean that many individual cells will only have a very small number of earthquakes or possibly none at all and so the N_{GR} are estimated on very small datasets. However, these initial N_{GR} values, (and also the initial b_{GR} values and maximum magnitudes), are modified using of an adaptive kernel smoothing method with Gaussian kernel function (Silverman, 1986; Stock and Smith, 2002). The smoothing procedure yields final values for the cell attributes, which account not only for the seismicity within the cell, but also account for the seismicity in all cells that are within a specified correlation distance of the cell. The correlation distance is assumed here to be 50 km (Stirling et al., 2002) and this ensures that the final smoothed attributes for each cell are based on larger datasets.

The adaptive kernel method is a five-step process that is applied separately for each of the three attributes. It is noted that the smoothing only takes place across a single grid depth layer and not between layers. The first step is to calculate a pilot estimate for each cell, which is a weighted sum of the initial value from the cell itself and the initial values from surrounding cells. For a given cell, the weight given in the summation to a value from a nearby cell *i* is calculated by the Gaussian kernel function given in Eq. 4.8, where d_{corr} is the correlation distance and d_{cell} is the distance between the two cells. Weights are only assigned to surrounding cells that are less than three times the correlation distance (150 km) from the cell being analysed. The pilot estimate is given by Eq. 4.9. The pilot estimate for each cell is then used to estimate a local bandwidth factor, b_f, for each cell, given by Eq. 4.10, where μ is the mean value across the entire grid (excluding zero values). The local bandwidth factor is used to calculate a revised set of weights for all cell pairs. First the distance limit for calculating weights is revised to 150 km multiplied by the local bandwidth factor, b_f. The revised weights are then calculated by Eq. 4.11. The final estimate for each cell is then given by Eq. 4.12.

$$weight_i = \exp\left(\frac{-d_{cell}^2}{d_{corr}^2}\right) \tag{4.8}$$

$$Pilot\,estimate = \frac{\sum\left[Initial\ value\right]_i \times weight_i}{\sum weight_i} \tag{4.9}$$

$$b_f = \sqrt{\frac{\mu}{Pilot\ estimate}} \tag{4.10}$$

$$weight_i^{'} = \exp\left(\frac{-d_{cell}^2}{d_{corr}^2 b_f^2}\right) \tag{4.11}$$

$$Final\,estimate = \frac{\sum\left[Initial\ value\right]_i \times weight_i^{'}}{\sum weight_i^{'}} \tag{4.12}$$

The a-value for each cell is calculated by Eq. 4.13, using the smoothed b_{GR} and N_{GR} and where $M_{\min 1} = 4$, $ct_1 = 2015 - 1964$, $M_{\min 2} = 5$, $ct_2 = 1964 - 1940$, $M_{\min 3} = 6.5$ and $ct_3 = 1940\text{-}1840$.

$$a = \log\left[\frac{N_{GR1} + N_{GR2} + N_{GR3}}{\left(ct_1 \times 10^{M_{\min 1} \times b}\right) + \left(ct_2 \times 10^{M_{\min 2} \times b}\right) + \left(ct_3 \times 10^{M_{\min 3} \times b}\right)}\right] \quad (4.13)$$

4.3.2 Stochastic event catalogue

For the seismic risk assessment Monte Carlo simulation, only earthquakes with $M_W \geq 5$ are considered at intervals of 0.1. This is slightly lower than the NSHM, which considers earthquakes with $M_W \geq 5.25$. For active faults, the return periods supplied by GNS Science relate only to the characteristic magnitude. Although active faults can generate smaller earthquakes, these are accounted for in the distributed seismicity sources. For the distributed seismicity sources, each cell is a potential source and any magnitude up to the maximum for that cell is possible. Therefore for active faults and distributed seismicity combined, there are a total of 61,204 possible magnitude/source combinations. For the active faults the annual rate of occurrence is defined in the dataset, and for distributed seismicity the annual rate of each magnitude/source combination can be determined by the Gutenberg-Richter relation from the calculated a-values and b-values. The total annual rate for earthquakes with $M_W \geq 5$ is 1.45 per year, which is equivalent to an earthquake occurring every 253 days. It is common for insurers conducting risk assessment to consider hazard-induced losses on an annual basis rather than a specific event basis and so it is proposed to generate a 10,000 year catalogue of events for this study. Since the annual event rate is 1.45, it is expected that a 10,000 year catalogue would generate more than 10,000 events and so the catalogue can be considered large in both year and event terms.

The generation of events is a two-step process. Firstly the time interval to the next event is determined in days and then a source and magnitude are assigned to the event. The time between events is modelled as a Poisson process, which is exponentially distributed. By

sampling a random number, U_1, from a uniform distribution between 0 and 1, the time to the next earthquake in years can be generated by Eq. 4.14, where λ is the rate parameter, which is 1.45 events per year. The magnitude and source of the event are determined by weighted random sampling from all magnitude/source combinations, with weights equal to the annual rate of each combination.

$$Time\ to\ event = \frac{\ln(U_1)}{\lambda} \tag{4.14}$$

Whilst for the precise geometry of events generated from active faults is known, for distributed seismic events, the determination of source only defines which cell and depth layer an earthquake occurs in. Many of the more recently developed GMPEs, such as the PEER-NGA equations, require more detailed geometries to be known and so additional attributes need to be estimated or assigned. The first attribute to be determined is where in the cell an earthquake is horizontally located, i.e. the surface projection of the top edge of rupture plane. It is assumed that the edges of the ruptures are straight lines and the centre point of the top edge is randomly assigned within the cell. The strike of the rupture is determined through a two-step process. Analysis of the 80 active faults in the model shows that their strikes are not evenly distributed. 19 faults have a strike less than 90°, 14 have a strike between 90° and 180°, 45 have a strike between 180° and 270° and just two faults have a strike between 270° and 360°. It is proposed to preserve these proportions within the distributed seismicity, so the first step to determine strike is to randomly assign a quadrant with weights according to the distribution observed amongst active faults. The second stage is to randomly sample a strike within the assigned quadrant assuming a uniform distribution. The length of the rupture, L_{rup}, is estimated using the empirical functions derived by Wells and Coppersmith (1994), which depend on faulting style as shown in Eq.'s 4.15 to 4.17. The equations predict a mean value for length and standard deviations for length are identified as 0.15 for strike-slip faults, 0.16 for reverse faults and 0.17 for normal faults. So that events of the same magnitude and fault type are not all assigned with

the same length, the standard deviations are used to add a random residual component to the predicted mean value, assuming a normal distribution with zero mean.

for strike-slip faults: $$L_{rup} = 10^{(-2.57+0.62M)} \quad (4.15)$$

for reverse faults: $$L_{rup} = 10^{(-2.42+0.58M)} \quad (4.16)$$

for normal faults: $$L_{rup} = 10^{(-1.88+0.5M)} \quad (4.17)$$

For most of the distributed sources zones, only one faulting style is assumed. However for sources in the Mid-Canterbury zone, both reverse and strike-slip fault mechanisms are possible. Analysis of the 20 active faults in the Mid-Canterbury zone shows that 18 of these are reverse faults and two are strike-slip faults. These proportions are preserved for distributed seismicity, so if a distributed seismic event is generated from the Mid-Canterbury zone, its faulting style is determined by weighted random sampling, with weights of 0.9 for reverse faults and 0.1 for strike-slip. Kaklamanos et al. (2011) propose guidelines for estimating unknown geometric input parameters for PEER-NGA equations. These guidelines are used to estimate dip, width and depth for distributed seismic events. Dip is assumed to be 50° for normal faulting events, 40° for reverse faulting events and 90° for strike-slip events. For width, they propose that the functions of Wells and Coppersmith (1994) are reasonable estimates, where the width, *W*, depends on faulting style as shown in Eq.'s 4.18 to 4.20.

for strike-slip faults: $$W = 10^{(-0.76+0.27M)} \quad (4.18)$$

for reverse faults: $$W = 10^{(-1.61+0.41M)} \quad (4.19)$$

for normal faults: $$W = 10^{(-1.14+0.35M)} \quad (4.20)$$

For depth, Kaklamanos et al. (2011) propose the method of Kaklamanos et al. (2010), as shown in Eq. 4.21, which determines the depth to the top of the rupture, Z_{TOR}, in terms of the dip, δ, and

hypocentral depth, Z_{HYP}, which can in turn be estimated from magnitude with the relationships proposed by Scherbaum et al. (2004) in Eq.'s 4.22 and 4.23.

$$Z_{TOR} = \max\left[\left(Z_{HYP} - 0.6W\sin\delta\right), 0\right] \tag{4.21}$$

for strike-slip faults: $$Z_{HYP} = 5.63 + 0.68M \tag{4.22}$$

for non-strike-slip faults: $$Z_{HYP} = 11.24 - 0.2M \tag{4.23}$$

The calculated geometric attributes allow the rupture plane to be fully described in three dimensions. It is also useful for GMPEs, to be able to describe the surface projection of the rupture plane. Since it is assumed that rupture planes are rectangular, their surface projection can be described by determining the latitude and longitude co-ordinates of each corner, shown in Figure 4.15, where (lat_0, $long_0$) are the co-ordinates of the centre point of the top edge of the rupture, (lat_1, $long_1$) are the co-ordinates of the corner along the strike direction and subsequent corners are numbered in a clockwise direction.

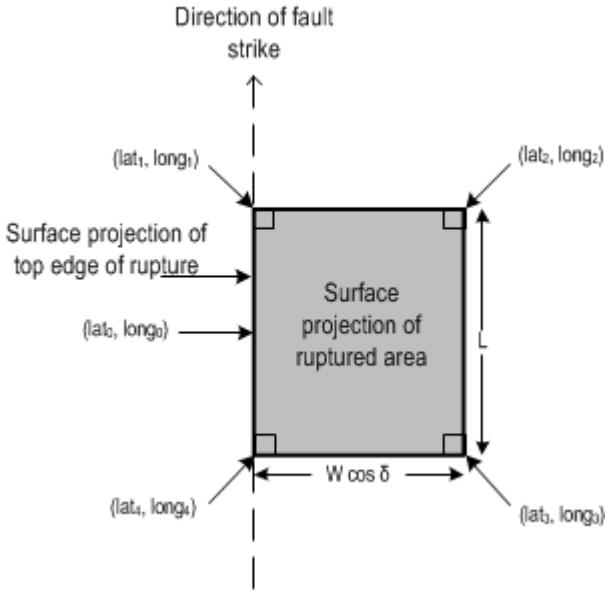

Figure 4.15 – Illustration of surface projection of rupture plane with geometric attributes

The haversine formula is used to measure distances and bearings along the Earth's surface (Sinnott, 1984). It can be rearranged so that given a pair of known co-ordinates (lat_a, $long_a$) for

some starting location, the co-ordinates of another location (lat_b, $long_b$) on the Earth's surface can be calculated by Eq.'s 4.24 and 4.25, if the bearing, *brng*, and distance, *dist*, between the two locations is known and the radius of the Earth, R_{earth}, is assumed to be 6,371 km. The inputs to the haversine formula for calculating the co-ordinates are listed in Table 4.8.

$$lat_b = \arcsin\left[\sin\left(lat_a\right)\cos\left(\frac{dist}{R_{earth}}\right)+\cos\left(lat_a\right)\sin\left(\frac{dist}{R_{earth}}\right)\cos\left(brng\right)\right] \tag{4.24}$$

$$long_b = long_a + \arctan 2\left[\cos\left(\frac{dist}{R_{earth}}\right)-\sin\left(lat_a\right)\sin\left(lat_b\right),\sin\left(brng\right)\sin\left(\frac{dist}{R_{earth}}\right)\cos\left(lat_a\right)\right] \tag{4.25}$$

Table 4.8 – Inputs to haversine formula to calculate co-ordinates of each corner of rupture surface projection

Corner	Start point	Distance =	Bearing =
1	0	$L_{rup}/2$	Strike
2	1	$W \cos \delta$	Strike + 90°, if Strike < 270° Strike – 270°, if Strike ≥ 270°
3	2	L_{rup}	Strike + 180, if Strike < 180° Strike – 180, if Strike ≥ 180°
4	0	$L_{rup}/2$	Strike + 180, if Strike < 180° Strike – 180, if Strike ≥ 180°

4.4 Step 3 – Hazard intensity prediction

For the calculation of hazard intensity values there are two possible approaches: either to calculate intensity at each site of interest (i.e. each infrastructure component) or to calculate intensity on a regularly-spaced grid or mesh. Although site-specific intensity is more precise, the grid/mesh approach is advantageous when the number of sites of interest is large, in particular due to the complexity of determining the covariance matrix for spatial correlation between sites, the computation time for which increases exponentially as the number of sites

increases (Weatherill et al., 2014) – in this study the electric power and water supply system contains nearly 7,000 elements between them. Furthermore, it is only when a site is close to an earthquake source that the difference between an intensity approximated from a nearby grid location is likely to differ noticeably from an intensity calculated based on an exact location. If necessary, the precision of the grid/mesh approach can be increased by increasing the resolution of the grid points/mesh cells. The use of a mesh or grid for calculating hazard intensity is common for infrastructure risk assessment platforms. Hazus (FEMA, 2015) adopts a grid with 100 points per 1° latitude and 1° longitude (Neighbors et al., 2013), which is approximately 0.9-1.1 km between points, depending on the precise latitude. SYNER-G adopts a variable mesh but with a minimum cell size of 1 km x 1 km (Franchin, 2014) and the RiskScape platform for New Zealand also adopts 1 km x 1 km cells as its highest resolution. However, it is known that commercial catastrophe models sometimes adopt higher resolutions for hazard analysis, particularly in areas of high seismicity (AIR, 2014). It is therefore proposed to test both a 1 km x 1 km mesh (294 cells) and a 500 m x 500 m mesh (1,007 cells) in this study at the validation stage in order to determine which should be adopted for the subsequent risk assessment. The two proposed meshes are shown in Figure 4.16.

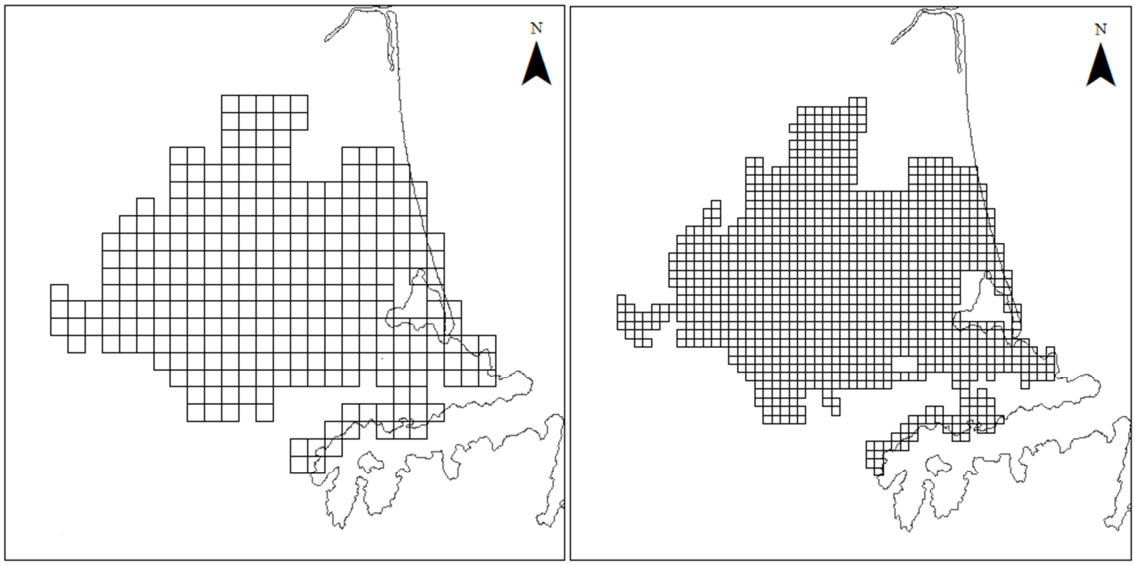

Figure 4.16 – 1km (left) and 500m (right) meshes for hazard analysis

4.4.1 Peak ground acceleration

Maximum horizontal PGAs are predicted in this study by the GMPE of Bradley (2013), which is a New Zealand-specific equation for shallow crustal earthquakes and therefore relevant to the seismic hazard risk for Christchurch. Bradley (2013) tested the applicability of five GMPEs – one New Zealand-specific and four foreign – by investigating the distributions of the residuals generated by each GMPE for accuracy and precision against observations from the New Zealand strong motion database (2,437 records from 213 events between 1973 and 2009). The analysis found that the most applicable was the foreign GMPE of Chiou et al., (2010), which is a modification of the PEER-NGA model of Chiou and Youngs (2008). However, the analysis was able to identify a number of discrepancies in this model, relating to dependence between the residuals and predictor variables. To rectify this, Bradley (2013) proposes five New Zealand-specific modifications to the Chiou et al. (2010) model: i) small magnitude scaling; ii) normal faulting factor; iii) site effects for site class A; iv) anelastic attenuation in the crust; and v) increased attenuation effects in the Taupo Volcanic Zone (North Island). Median predictions and confidence intervals from the resulting GMPE were compared to PGA observations from the Darfield and Christchurch earthquakes and found to achieve sufficient precision and accuracy, whilst also demonstrating no dependence between the residuals and predictor variables. The final Bradley (2013) GMPE model for median PGA at a site (centre of a cell) is shown in Eq.'s 4.26 to 4.29, excluding the term for the Taupo Volcanic Zone which is not relevant for Christchurch. The predictor variables for the model are: M_W, moment magnitude; R_{rup}, the slant distance to the closest point on fault rupture plane (km); R_{JB}, Joyner-Boore distance, which is the horizontal distance to the surface projection of the fault rupture plane (km); R_X, distance (km) from the surface projection of the up-dip edge of the fault plane, measured perpendicular to the fault strike (positive in the down-dip direction); F_{HW}, hanging-wall flag (1 for $R_X \geq 0$ and 0 for $R_X < 0$); δ, the fault dip angle; Z_{TOR}, depth to top of the fault rupture plane; F_{RV}, reverse faulting flag; F_{NM}, normal faulting flag; V_{S30}; average shear-wave velocity for the top 30 m of the site (m/s); and $Z_{1.0}$, depth to shear-wave velocity of 1.0 km/s

(m). Most of these values can be determined straightforwardly from the physical and geometric attributes defined for each event at the hazard generation stage. V_{S30} is obtained from the USGS Global Server (USGS, 2013), which provides global estimates on a grid with spacing of approximately 674 m x 674 m. At any location in the study area, V_{S30} is assumed to be the same as the value at the grid point closest to the site. Chiou and Youngs (2008) recommend estimating a median value of $Z_{1.0}$ based on V_{S30} as shown in Eq. 4.30.

$$\begin{aligned}\ln\left(PGA_{ref}\right) = &-1.1985 + \left[0.1F_{RV} - 0.455F_{NM} + 0.0512\left(\min\left\{Z_{TOR},10\right\} - 4\right)\right] \\ &+ 1.06\left(M_W - 6\right) - 0.1469\ln\left(1 + e^{2.996\left(5.85 - M_w\right)}\right) \\ &- 2.1\ln\left[R_{rup} + 6.16\cosh\left\{0.4893\max\left(M - 3, 0\right)\right\}\right] \\ &+ 1.6\ln\left(\sqrt{R_{rup}^2 + 2500}\right) \\ &- \left\{0.0096 + \frac{0.0048}{\cosh\left[\max\left(M - 4, 0\right)\right]}\right\} R_{rup} \\ &+ 0.79F_{HW}\tanh\left(\frac{R_X\cos^2\delta}{1.5005}\right)\left\{1 - \frac{\sqrt{R_{JB}^2 + Z_{TOR}^2}}{R_{rup} + 0.001}\right\}\end{aligned} \tag{4.26}$$

and

$$\begin{aligned}\ln\left(PGA\right) = \ln\left(PGA_{ref}\right) &- 0.4417\ln\left(\frac{\min\left(V_{S30}, V_1\right)}{1130}\right) + b_{BRAD}\ln\left(\frac{PGA_{ref} + 0.102151}{0.102151}\right) \\ &+ 0.2289\left(1 - \frac{1}{\cosh\left[0.014996\max\left(0, Z_{1.0} - 580\right)\right]}\right) \\ &+ \frac{0.07}{\cosh\left[0.15\max\left(0, Z_{1.0} - 15\right)\right]}\end{aligned} \tag{4.27}$$

where

$$V_1 = \min\left(1130\max\left\{\left(\frac{T}{0.75}\right)^{-0.11}, 1\right\}, 1800\right) \tag{4.28}$$

and

$$b_{BRAD} = -0.1417\left\{e^{-0.00701\left(\min\left(V_{S30},1130\right)-360\right)} - e^{-0.00701\left(1130-360\right)}\right\} \tag{4.29}$$

$$Z_{1.0} = \exp\left[28.5 - \frac{3.82}{8}\ln\left({V_{S30}}^{8} + 378.7^{8}\right)\right] \tag{4.30}$$

Visual representations of the three distance measurements are shown in Figure 4.17 and Figure 4.18. A new measurement, source-to-site azimuth, α, is introduced. The source-to-site azimuth can be thought of as a bearing between the closest point on the surface projection of top edge of the rupture and the site, and is measured relative to the strike rather than to north. Sites perpendicular to the edge have an azimuth of 90° or -90°. Sites on the hanging wall side of a rupture have positive azimuths (0° to 180°) and sites on the footwall side of a rupture have anti-clockwise azimuths (-180° to 0°). R_{JB} between a site and the surface projection is measured automatically using the 'geosphere' package for R (Hijmans et al., 2015).

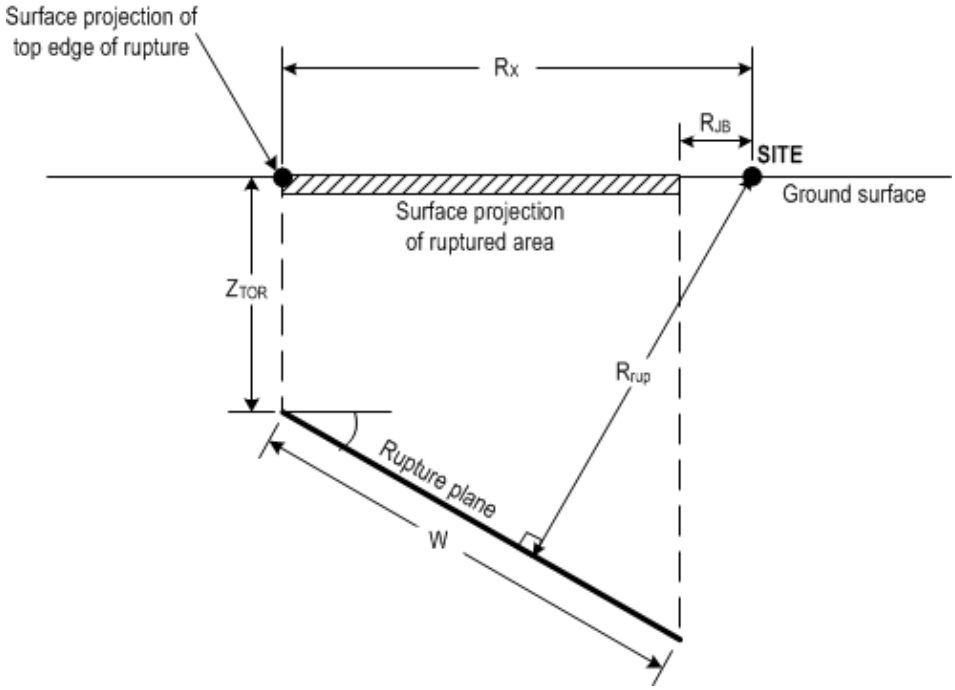

Figure 4.17 – Illustration depicting the three distance measurements needed for predicting PGA (redrawn from Kaklamanos et al. (2011))

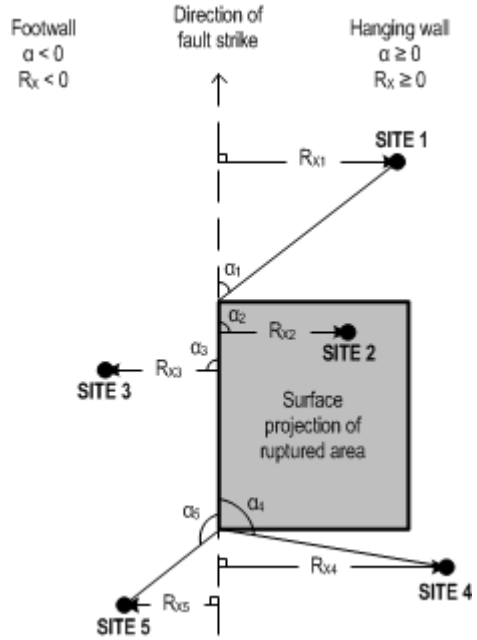

Figure 4.18 – Illustration depicting the measurement of source-to-site azimuth, α, and distance, R_X, relative to the surface projection of the rupture (redrawn from Kaklamamos et al. (2011))

R_X is positive for sites on the hanging wall side and negative for sites on the footwall side Measurement of R_X depends on the location of the site relative to the top edge of the rupture. Sites can be categorized into one of nine cases depending on the value of α and, if the site is on the hanging wall side, whether the perpendicular distance between the strike and the site (R_{JB} |tan α|) is less than or greater than the perpendicular distance between the top and bottom edges of the rupture plane (W cos δ). An illustration of the nine cases is shown in Figure 4.19. For cases 4, 5 and 6, R_X can be measured automatically by the 'geosphere' package, since it is simply the perpendicular distance to the top edge of the rupture. For other cases however, R_X has to be determined analytically based on surface trigonometry. The relationships, derived by Kaklamanos et al. (2011), are given in Eq.'s 4.31 to 4.33.

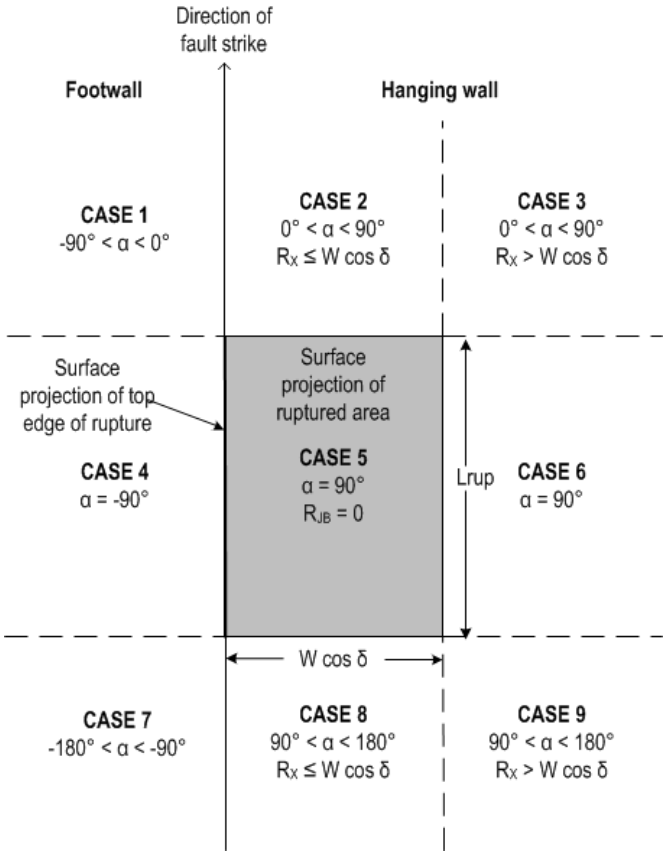

Figure 4.19 – Illustration of nine cases for combinations of source-to-site azimuth, α, and distance R_x (redrawn from Kaklamanos et al. (2011))

for cases 1 and 7: $$R_X = R_{JB} \sin\alpha \tag{4.31}$$

for cases 2 and 8: $$R_X = R_{JB} \left|\tan\alpha\right| \tag{4.32}$$

for cases 3 and 9: $$R_X = R_{JB} \tan\alpha \cos\left[\alpha - \sin^{-1}\left(\frac{W\cos\delta\cos\alpha}{R_{JB}}\right)\right] \tag{4.33}$$

R_{rup} is calculated from Eq.'s 4.34 to 4.41, which introduce two new interim measurements: R'_{rup}, the in-plane rupture distance, and R_Y, the closest distance to the rupture surface projection measured parallel to the strike.

for δ = 90°: $$R_{rup} = \sqrt{R_{JB}^2 + Z_{TOR}^2} \tag{4.34}$$

otherwise: $$R_{rup} = \sqrt{\left(R'_{rup}\right)^2 + R_Y^2} \tag{4.35}$$

where:

for $R_X < Z_{TOR} \tan\delta$: $$R'_{rup} = \sqrt{R_X^2 + Z_{TOR}^2} \tag{4.36}$$

for $Z_{TOR} \tan\delta \le R_X \le Z_{TOR} \tan\delta + W \sec\delta$: $$R'_{rup} = R_X \sin\delta + Z_{TOR} \cos\delta \tag{4.37}$$

for $R_X > Z_{TOR} \tan\delta + W \sec\delta$: $$R'_{rup} = \sqrt{\left(R_X - W\cos\delta\right)^2 + \left(Z_{TOR} + W\sin\delta\right)^2} \tag{4.38}$$

and:

for α = ±90°: $$R_Y = 0 \tag{4.39}$$

for α = 0° or ±180°: $$R_Y = R_{JB} \tag{4.40}$$

otherwise: $$R_Y = \left|R_X \cot\alpha\right| \tag{4.41}$$

Bradley (2013) adopts the Chiou and Youngs (2008) GMPE standard deviation model without change. The model for the inter-event residual standard error, τ, is shown in Eq. 4.42 and the model for the intra-event residual standard error, σ, is shown in Eq. 4.43, where $F_{Inferred}$ is 1 if V_{S30} is inferred from geology or 0 otherwise; $F_{Measured}$ is 1 if V_{S30} is measured or 0 otherwise.

$$\tau = 0.3437 - 0.04\left[\min\left\{\max\left(M,5\right),7\right\} - 5\right] \tag{4.42}$$

$$\sigma = \left[0.4458 - 0.05\left(\min\left\{\max\left(M,5\right),7\right\} - 5\right)\right] \times \sqrt{\left(0.8F_{Inferred} + 0.7F_{Measured}\right) + \left(1 + NL\right)^2} \tag{4.43}$$

where:

$$NL = \left(b\frac{PGA_{ref}\, e^{\eta}}{PGA_{ref}\, e^{\eta} + 0.102151}\right) \tag{4.44}$$

The non-linear term, *NL*, accounts for the non-linear response in the intra-event variability and the parameter η is a random effects term which should be determined by integration. However Chiou and Youngs (2008) propose a Taylor series approximation, shown in Eq. 4.45, which allows the total standard error, σ_T, to be calculated without requiring integration.

$$\sigma_T^2 = \left(1 + NL_0\right)^2 \tau^2 + \sigma_{NL_0}^2 \tag{4.45}$$

In Eq. 4.45, NL_0 is the value of *NL* calculated for the case where η is 0 and σ_{NLo} is the value of σ calculated when η is 0.

4.4.2 Spatial correlation

It is proposed to develop a new empirical Christchurch-specific spatial correlation model based on data recorded at strong motion stations during the Canterbury earthquake sequence and using the method proposed by Goovaerts (1997) as described in section 2.3.3. There are 33 strong motion stations in the Christchurch City Council area, shown in Figure 4.20, and recordings have been obtained from GeoNet (GNS Science, 2014) for four earthquakes: the Darfield earthquake, the Christchurch earthquake, an M_W 5.9 earthquake on 13th June 2011 and an M_W 5.8 earthquake on 23rd December 2011. However, not all stations recorded observations for all earthquakes. There are a total of 86 PGA observations across the four earthquakes, creating 1,796 correlation observations for the empirical model. For PGA, spatial correlation applies to the intra-event residual, ε, which is randomly generated across sites for a given event. Here it is proposed to use the normalised intra-event residual, $\varepsilon' = \varepsilon/\sigma$, as the variable to be correlated. For a site-separation distance, *h*, the observed stationary semivariogram is derived by calculating the average observed value of γ for all N_{SC} pairs of sites separated by *h*, as shown in Eq. 4.46.

$$\hat{\gamma} = \frac{1}{2N} \sum_{i=1}^{N_{SC}} \left\{ \varepsilon'_{u_i} - \varepsilon'_{u_i+h} \right\}^2 \tag{4.46}$$

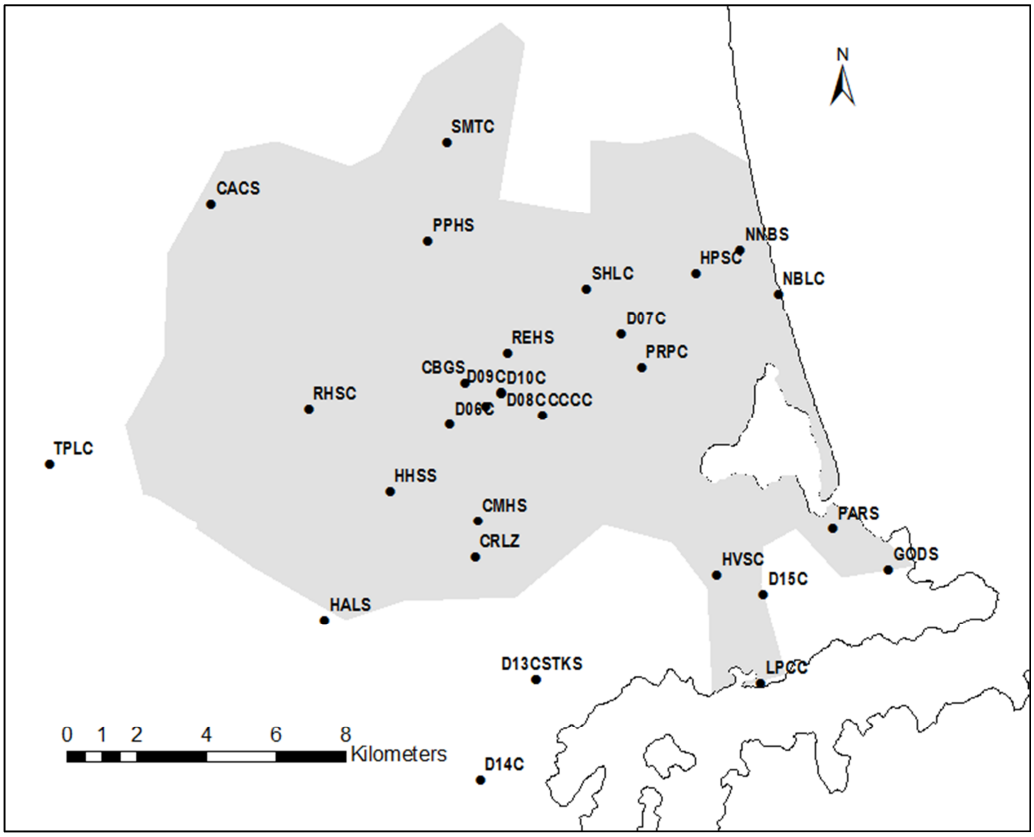

Figure 4.20 – Locations of 31 strong motions station sites for spatial correlation analysis (the remaining two sites are located on the Banks Peninsula, 10km to the south of the city and are not shown here for clarity)

Since this is an empirical analysis, ε' is defined as the normalised difference between the observed PGA and the PGA predicted by the Bradley (2013) GMPE at a site. This is repeated for all values of h to give an observed semivariogram plot to which a continuous semi-definite (non-negative) function is fit, so that the semivariogram and correlation can be predicted for any value of h. To confirm that a selected GMPE is adequate for empirical analysis of the correlation structure of ground motions, Goda and Atkinson (2009) propose inspection of the plot of normalised intra-event residuals, ε', against distance to rupture R_{rup}, as shown in Figure 4.21. This shows that 67 of the 86 residuals (78%) are within one standard deviation of the predicted median and that the residuals show negligible dependence on distance to rupture, thus confirming that the Bradley (2013) GMPE is adequate for analysing the correlation structure. The distribution of site separation distances, h, amongst the 1,796 correlation observations is shown in Figure 4.22.

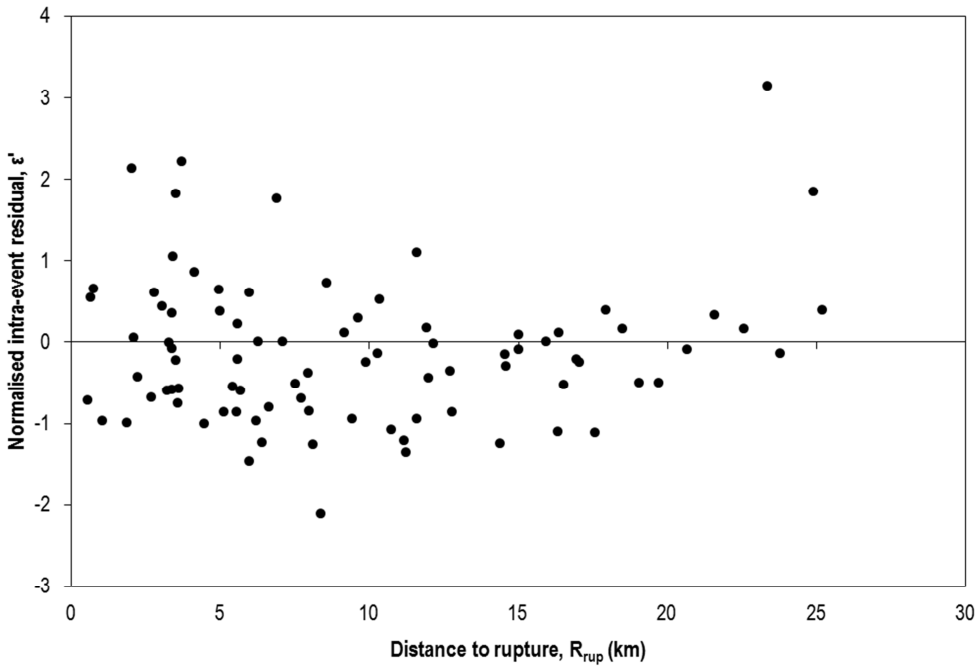

Figure 4.21 – Plot of normalised intra-event residuals against distance to rupture as calculated by Bradley (2013) GMPE

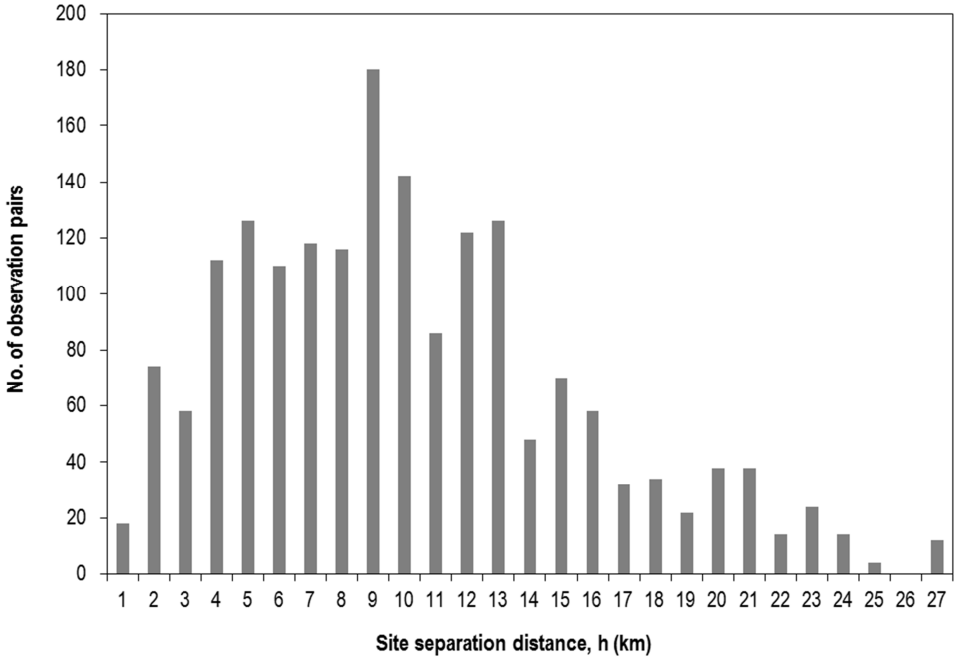

Figure 4.22 – Distribution of separation distances amongst the 1,796 site pair observations from Christchurch

However, when creating the observed semivariogram, a balance has to be achieved between ensuring there are sufficient points with which to fit the model and ensuring that each point is itself based on a sufficient number of observations to be a reliable estimate of γ for that value of h. Esposito and Iervolino (2011) suggest that each point should be based on a minimum sample size of 30. Therefore it is necessary to put the data into bins for h. The maximum separation distance in the dataset is 26.7 km and so the proposed bin widths are 1 km, 2 km, 3 km and 4 km. Whilst a smaller bin width generates more data points, a larger bin width generates data points based on larger sample sizes. Accuracy in the spatial correlation model is most important at shorter separation distances and a comparison of the two bin widths shows that up to 12 m separation (an arbitrary cut-off chosen for convenience since it is divisible by 2, 3 and 4), the 1 km bin width yields twelve points with an average sample size of 105; the 2 km bin width yields six data points with an average sample size of 210; the 3 km bin width yields four data points with an average sample size of 316; and the 4 km bin width yields three points with an average sample size of 421.

The observed semi-variogram plots are shown in Figure 4.23 for data points based on samples larger than 30 as suggested by Esposito and Iervolino (2011). The plot for the 1km bins exhibits a very wide scatter across all distances and as so this dataset is not considered for further analysis. The other plots show the expected general trend that the semi-variogram increases as site separation distance increases. In all three cases however, the data points at the highest separation distances show a pronounced decrease in the semi-variogram value compared to the trend. It is notable that for the 2 km bin width, the two rightmost data points are based on just 52 and 38 samples, whilst for the 3 km bin width, the rightmost data point is based on just 52 samples. These are the three smallest samples of all the points plotted, and highlight the importance of ensuring that data points are based on sufficiently large samples.

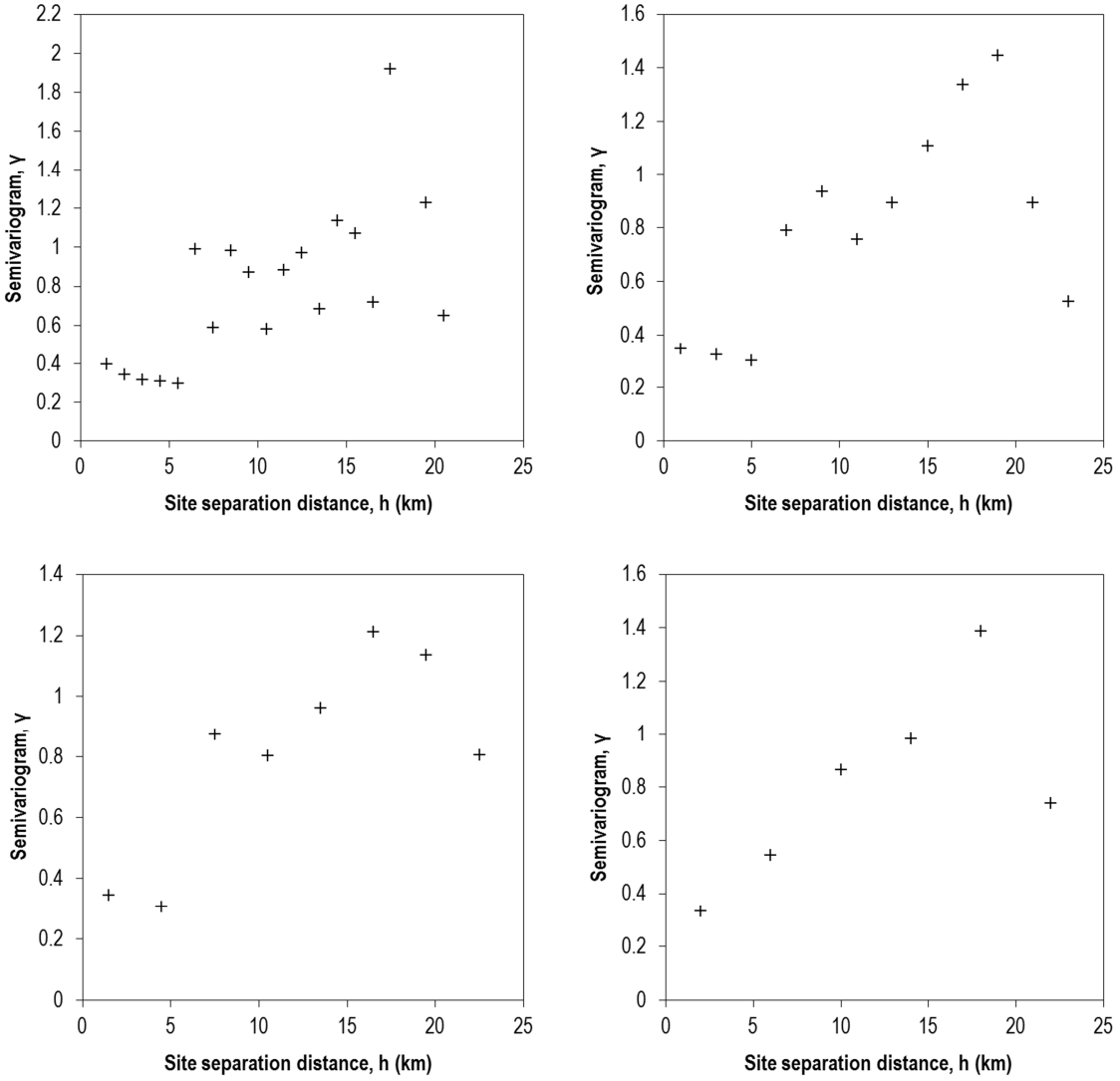

Figure 4.23 – Observed semi-variogram plots for different separation distance bin widths:1km (top left), 2km (top right), 3km (bottom left), 4km (bottom right)

To fit a model to the observed semivariogram, a functional form has to be assumed that fits the conditions of $\gamma = 0$ for $h = 0$ and $\gamma = 1$ for $h = \infty$. Five of the functional forms discussed in the literature review (see Table 2.3) are tested in this study. The models by Boore et al. (2003) and Wang and Takada (2005) are single parameter models fit to an optimum value of α_B and α_{WT} respectively. The model by Goda and Atkinson (2009) is a three-parameter model, seeking to find the optimum values of α_{GA}, β_{GA} and κ_{GA}. No physical meaning is defined for any of the parameters in these three models. Since the variance of the normalised intra-event residual is 1, in this study the three forms proposed by Jayaram and Baker (2009) are all effectively single parameter models, seeking the optimum value of b_{JB}, which is known as the range and has a

physical meaning. For the exponential and Gaussian models, the range is the separation distance at which γ is 0.95 times the sill, i.e. at which $\gamma = 0.95$ since $a_{SC} = 1$. For the spherical model, the range is the separation distance at which γ equals the sill, i.e. at which $\gamma = 1$. In Table 2.3 it can be seen that the exponential model proposed by Jayaram and Baker (2009) is effectively the same form as the Wang and Takada (2005) model since α_{WT} is equivalent to $3/b_{JB}$. So that the exponential model can be compared to the other models proposed by Jayaram and Baker (2009), it is retained in the model testing and the Wang and Takada (2005) model is excluded.

To determine the optimum values of the parameters for each functional form, the least squares method is used to minimise the sum of the squared residuals (SSR) between the predicted and observed data points. However, as discussed by Jayaram and Baker (2009), the focus of the optimisation should be to ensure that the model fits well at short separation distances when the correlation is high. At greater distances, when the correlation is low, the effect that a site has on the residual of another site is likely to be negligible. Jayaram and Baker (2009) do not explicitly define a maximum value that can be considered to be a short separation distance. However, a visual inspection of their empirical semivariogram models for the 1994 M_W 6.7 Northridge earthquake in Los Angeles and the 1999 M_W 7.6 Chi-Chi earthquake in Taiwan, indicate that the model is fit to optimize performance for site separation distances up to 15 km. It is therefore proposed to use a 15 km cut-off for application of the least squares method for model fitting.

The optimisation bias towards short separation distances means that for 4 km bins, the model is fitted to just four data points, compared to five data points for 3 km bins and seven data points for 2 km bins. It is therefore proposed to consider only the 2 km and 3 km bin widths for model fitting. The optimised parameters for both bin widths for all of the models is shown in Table 4.9. It is notable that with the exception of the Goda and Atkinson (2009) model, for each model the optimised parameters are very similar irrespective of the bin width used. When a 3 km bin is used, the exponential and spherical models of Jayaram and Baker (2009) fit short separation distances equally well. When a 2 km bin is used, the spherical model is the best fitting model. SSRs of the 3 km bin models and the 2 km bin models cannot be compared directly since the 2

km bin analysis contains more data points. For the exponential and spherical models derived from 3 km bin data, the average squared residual per observation is 0.019. For the spherical model derived from the 2 km bin data, the average squared residual per observation is 0.022. This indicates that the models fit to the 3 km bin data are a slightly better fit than the model fit to the 2 km bin data.

Table 4.9 – Optimised parameters for semivariogram models fitted to Christchurch data

Model	**3km bin width**		**2km bin width**	
	Parameters	**SSR (h < 15km)**	**Parameters**	**SSR (h < 15km)**
Boore et al. (2003)	$\alpha_B = 0.44$	0.153	$\alpha_B = 0.45$	0.292
Goda and Atkinson (2009)	$\alpha_{GA} = 0.04$, $\beta_{GA} = 0.64$, $\kappa_{GA} = 5$	0.107	$\alpha_{GA} = 0.03$, $\beta_{GA} = 0.93$, $\kappa_{GA} = 5$	0.193
Jayaram and Baker (2009) (exponential)	$b_{JB} = 18$	0.095	$b_{JB} = 17.5$	0.189
Jayaram and Baker (2009) (Gaussian)	$b_{JB} = 11$	0.125	$b_{JB} = 11$	0.215
Jayaram and Baker (2009) (spherical)	$b_{JB} = 14$	0.095	$b_{JB} = 15$	0.173

These two models are plotted in Figure 4.24 alongside the observations for the 3 km bins. The primary difference between the two models is that the spherical model has a separation distance cut-off point, equal to the range, b_{JB}, above which the semi-variogram equals 1 and hence the correlation is 0, whereas the exponential model continues to approach 1 as separation distances increases. The evidence to support the selection of either the exponential or spherical over the other is inconclusive. It is proposed to adopt the exponential model for this case study, since it yields a closer fit to the observation from the bin with the shortest site separation distance (<3 km). For simulating fields of spatially correlated intra-event residuals using this model, it is proposed to use the method of Weatherill et al. (2014), described in section 2.3.3.

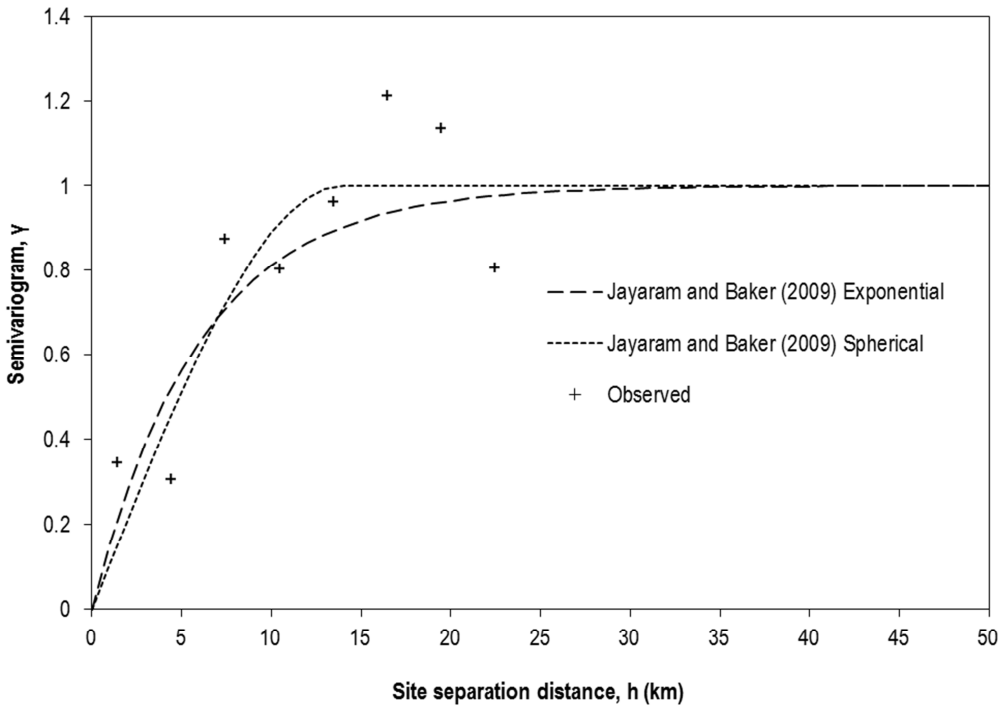

Figure 4.24 – Plots of the two semivariogram models that fit best to the observed data when optimised to minimise errors at separation distances up to 15km

4.4.3 Peak ground velocity

There are no GMPEs for maximum horizontal PGV that are specific to New Zealand. However, the GMPE used here for PGA (Bradley, 2013) is based on the PEER-NGA models by Chiou and Youngs (2008) and Chiou et al. (2010). The NGA models do include equations for predicting PGV and so to be as consistent as possible with PGA prediction, it is proposed to use these models here. The Chiou and Youngs (2008) model is proposed for earthquakes with M > 5.5 and the Chiou et al. (2010) model is proposed for earthquakes with M ≤ 5.5. Both models have the same functional form, but with variation in the coefficients c_1, c_3, c_m and $c_{\gamma 2}$. The different values for these coefficients used in each model are shown in Table 4.10 and the base model is shown in Eq.'s 4.47 and 4.48, excluding the aftershock flag term, which is not required in this research since the hazard model only predicts main shock occurrences.

Table 4.10 – Coefficients for GMPEs for PGV prediction

Coefficient	Chiou et al. (2010) M ≤ 5.5	Chiou and Youngs (2008) M > 5.5
c_1	2.2877	2.2884
c_3	1.7267	3.45
$c_{\rm m}$	5.49	4.2979
$c_{\gamma 2}$	-0.000687	-0.00625

$$\begin{aligned}\ln\left(PGV_{ref}\right) = c_1 &+ \left[0.1094F_{RV} - 0.0626F_{NM} + 0.0207\left(Z_{TOR} - 4\right)\right] \\ &+ 1.06\left(M - 6\right) - \frac{1.06 - c_3}{1.06}\ln\left(1 + e^{1.648\left(c_m - M\right)}\right) \\ &- 2.1\ln\left[R_{rup} + 5.17\cos\left\{0.4407\max\left(M - 3, 0\right)\right\}\right] \\ &+ 1.6\ln\left(\sqrt{R_{rup}^2 + 2500}\right) \\ &- \left\{0.00275 - \frac{c_{\gamma 2}}{\cosh\left[\max\left(M - 4, 0\right)\right]}\right\}R_{rup} \\ &+ 0.3079F_{HW}\tanh\left(\frac{R_X\cos^2\delta}{2.669}\right)\left\{1 - \frac{\sqrt{R_{JB}^2 + Z_{TOR}^2}}{R_{rup} + 0.001}\right\}\end{aligned} \tag{4.47}$$

and

$$\begin{aligned}\ln\left(PGV\right) = \ln\left(PGV_{ref}\right) &- 0.7861\min\left(\ln\left(\frac{V_{S30}}{1130}\right), 0\right) \\ &- 0.0699\left\{e^{-0.008444\left(\min\left(V_{S30}, 1130\right) - 360\right)} - e^{-0.008444\left(1130 - 360\right)}\right\}\ln\left(\frac{PGV_{ref} + 5.41}{5.41}\right) \\ &+ 0.2899\left(1 - \frac{1}{\cosh\left[0.006718\max\left(0, Z_{1.0} - 459\right)\right]}\right) \\ &+ \frac{0.1138}{\cosh\left[0.15\max\left(0, Z_{1.0} - 15\right)\right]}\end{aligned} \tag{4.48}$$

Both models adopt the same function for the total residual standard error as used for the PGA GMPE but with changes to the coefficient values used to calculate τ, σ and NL as shown in Eq.'s 4.49 to 4.51.

$$\tau = 0.2539 - 0.0089\left[\min\left\{\max\left(M,5\right),7\right\}-5\right] \tag{4.49}$$

$$\sigma = \left[0.4496 - 0.0471\left(\min\left\{\max\left(M,5\right),7\right\}-5\right)\right] \times\sqrt{\left(0.7504F_{Inferred}+0.7F_{Measured}\right)+\left(1+NL\right)^2} \tag{4.50}$$

where:

$$NL = \left(b\frac{PGV_{ref}e^{\eta}}{PGV_{ref}e^{\eta}+5.41}\right) \tag{4.51}$$

4.4.4 Ground shaking cross-correlation

Correlation between PGA and PGV intra-event residuals is accounted for in ground motion simulation using the sequential conditional simulation method proposed by Weatherill et al. (2014), which assumes that the correlation between residuals of PGA and PGV, $\rho_{PGA,PGV}$ is known. It is proposed to estimate the correlation here by calculating the Pearson product-moment correlation coefficient of the PGA and PGV observations from the same strong motion station records used in the spatial correlation analysis. The correlated total residual for PGV at a site is randomly sampled from a normal distribution, with a conditional mean, PGV_{cond} and a conditional standard deviation, $\sigma_{PGVcond}$, calculated by Eq.'s 4.52 and 4.53, respectively.

$$\ln\left(PGV_{cond}\right) = \ln\left(PGV_{med}\right) + \rho_{PGA,PGV}\sigma_{T,PGV}\frac{\varepsilon_{T,PGA}}{\sigma_{T,PGA}} \tag{4.52}$$

$$\sigma_{T,PGVcond} = \sigma_{T,PGV}\sqrt{1-\rho^2_{PGA,PGV}} \tag{4.53}$$

4.4.5 Liquefaction susceptibility

Both Brackley (2012) and CERA (2013) have produced liquefaction risk maps for different parts of Christchurch and its surrounding districts, using a scale consisting of four levels of risk. These maps have been based on topographic, geologic, soil type and borehole data and they

have been combined by Environment Canterbury into a single liquefaction susceptibility map shown in Figure 4.25 (ECan, 2014). The four zones are: 'None', 'Low', 'Moderate' and 'High'.

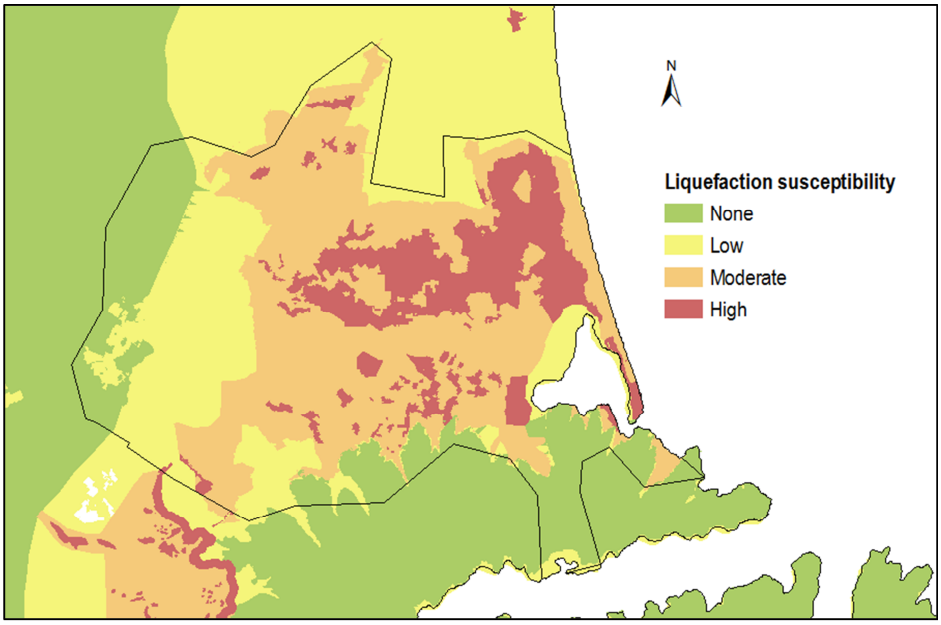

Figure 4.25 – Liquefaction susceptibility map for Christchurch and surrounding districts, based on data from ECan (2014)

4.4.6 Liquefaction triggering

It is proposed to use observations from the Darfield and Christchurch earthquakes to test the accuracy of the three simplified methods for predicting liquefaction triggering described in section 2.3.4: the shear-wave velocity method for Liquefaction Potential Index (LPI) (Iwasaki et al., 1984; Andrus and Stokoe, 2000); Hazus (FEMA, 2015) and Zhu et al. (2015). The following sections describe how the methods are applied to Christchurch and provide a summary and interpretation of the prediction results and diagnostic scores.

4.4.6.1 Application of methods

To implement the LPI method for Christchurch, the depth to ground water is assumed to be 2 m across Christchurch, reflecting the averages described by Giovinazzi et al. (2011) (0-2 m in the eastern suburbs and 2-3 m in the western suburbs) and soil unit weights are assumed to be 17 kPa above the water table and 19.5 kPa below the water table, as suggested by Wotherspoon et

al. (2014). Wood et al. (2011) have published V_S profiles for 13 sites across Christchurch obtained using surface wave testing methods. In GIS, the profiles are converted to point data for each 1 m depth increment from 0-20 m, so that each point represents the V_S at that site for a single soil layer and there are a total of 13 points for each soil layer. Ordinary kriging (with log transformation to ensure non-negativity) is applied to the points in each soil layer to create interpolated V_S raster surfaces. Whilst Andrus and Stokoe (2000) advise that the maximum V_{S1} (the stress-corrected shear wave velocity) can range from 200-215 m/s depending on fines content, subsequent work by Zhou and Chen (2007) indicates that the maximum V_{S1} could range between 200-230 m/s. In the absence of specific fines content data, a median value of 215 m/s is assumed. In practice, a soil layer may have a value of V_{S1} that is below this threshold but not be liquefiable because the soil is not predominantly clean sand. Due to the regional scale of this analysis though, site-specific soil profiles (as distinct from V_S profile) are not taken into account in determining whether a soil layer is liquefiable. Goda et al. (2011) suggest the use of 'typical' soil profiles to determine the liquefaction susceptibility of a soil layer at a regional scale. Borehole data at sites close to the 13 V_S profile sites are available from the New Zealand Geotechnical Database (EQC, 2016). These indicate that in the eastern suburbs of Christchurch, soil typically consists of clean sand to 20 m depth, with some layers of silty sand. On the western side of Christchurch however there is an increasing mix of sand, silt and gravel in soil profiles, particularly at depths down to 10 m. Therefore it is possible, particularly in western suburbs, that the calculated V_{S1} values may indicate liquefiable soil layers when they are in fact not, which could lead to overestimation of LPI and the extent of liquefaction. For testing, this model is referred to as LPI1.

One of the critical considerations for insurers is availability of model input data. Although in this case study the use of V_S, for which profiles exist in the literature, negates the requirement for ground investigation, for most sites V_S data are not commonly in the public domain. As is the case with Christchurch, even when data are available, it is not necessarily across an entire study area, thus requiring geostatistical techniques to interpolate. Consequently, although this

method is applicable here, it may not be in other study areas. Giving consideration to the wider needs of the insurance sector, two approaches are proposed to extend the applicability of the LPI model by approximating V_S from more readily available datasets.

The first approach, referred to as model LPI2, uses V_{S30}, the average shear wave velocity across the top 30 m of soil, as a constant proxy for V_S for all soil layers in 2.17. Global point estimates for V_{S30} at approximately 674 m grid intervals can be found on the web-based USGS Global V_{S30} Map Server (USGS, 2013). Hence the use of V_{S30} fulfils the requirement for easily accessible data. The disadvantage of this approach is that the likelihood of liquefaction occurrence in the LPI method is controlled by the presence of soil layers near the surface with low V_S. Furthermore, there is a maximum value of V_S at which liquefaction can occur. Hence the use of V_{S30} as a proxy for all layers will result in an overestimation of V_S, CRR and FS^* at layers closer to the surface, and therefore an underestimation of LPI and liquefaction risk.

The second approach, referred to as model LPI3, proposes the use of the same V_{S30} data but manipulates it to simulate a more realistic V_S profile in which velocities decrease towards the surface rather than being constant. Boore (2004) proposes empirical functions to extrapolate V_{S30} values in situations where shear wave velocity data are only known up to shallower depths, based on observations from the United States and Japan. It is proposed to use these empirical functions in reverse – to back-calculate shallower average shear wave velocities from V_{S30} data from the USGS Global Server (USGS, 2013). For simplicity it is proposed to only use the empirical functions to calculate $V_{S(0\text{-}10)}$ (average shear wave velocity across top 10 m) and $V_{S(10\text{-}20)}$ (average shear wave velocity across top 20 m). The calculated value for $V_{S(0\text{-}10)}$ can then be used as a proxy for V_S at all soil layers between 0-10 m depth and both the $V_{S(0\text{-}10)}$ and $V_{S(10\text{-}20)}$ values can be used to determine an equivalent proxy for all soil layers between 10-20 m. From manipulation of the Boore (2004) empirical functions and the formula for calculating averaged shear wave velocities, Eq.'s 4.54 and 4.55 determine the proxies to be used in the two depth ranges.

$$V_{S(0-10)} = 10^{\left(\frac{\log V_{S30} - 0.042062}{1.0292}\right)} \tag{4.54}$$

$$V_{S(10-20)} = 2 \times 10^{\left(\frac{\log V_{S30} - 0.025439}{1.0095}\right)} - V_{S(0-10)} \tag{4.55}$$

For the Hazus model, since it uses a different liquefaction susceptibility scale to the one that has been developed for Christchurch in Figure 4.25, it is necessary to map the four Christchurch zones onto four of the Hazus zones. The simplest way is to assume that 'Low' and 'Very low' in Hazus are equivalent to 'Low' in Christchurch, and that 'High' and 'Very high' in Hazus are equivalent to 'High' in Christchurch. However this creates problems when trying to selecting parameters from Table 2.4. Three approaches for selecting parameters are proposed and these are described in Table 4.11. The accuracy of all three approaches are tested against observations in this analysis and are given corresponding model names: HAZ1, HAZ2, HAZ3.

Table 4.11 – Conversion between Canterbury and Hazus liquefaction susceptibility zones for three implementations of Hazus methodology

Christchurch liquefaction susceptibility zones	**Equivalent Hazus liquefaction susceptibility zone parameters**		
	HAZ1	**HAZ2**	**HAZ3**
None	None	None	None
Low	Low	Very low	Average low and very low
Moderate	Moderate	Moderate	Moderate
High	High	Very high	Average high and very high

For the Zhu et al. (2015) all three models are tested and the global, regional and local models are given the names ZHU1, ZHU2 and ZHU3, respectively. All input data are available from USGS websites (USGS, 2013; 2014), although here the PGA map from the New Zealand Geotechnical Database (New Zealand Earthquake Commission, 2016) is used as a substitute for USGS ShakeMaps. A list of all models is shown in Table 4.12.

Table 4.12 – Liquefaction prediction models tested in this analysis

Model	Description
LPI1	LPI with known V_S profiles
LPI2	LPI with V_{S30} as V_S proxy
LPI3	LPI with simulated V_S profiles
LPIref	LPI calculated from SPT results
HAZ1	Hazus with 'direct' conversion of susceptibility zones
HAZ2	Hazus with 'extreme' susceptibility zones
HAZ3	Hazus with 'average' conversion of susceptibility
ZHU1	Global model by Zhu et al. (2015)
ZHU2	Regional model by Zhu et al. (2015)
ZHU3	Local model by Zhu et al. (2015)

Table 4.12 includes an additional model LPIref. LPIref is a dataset of LPI values measured after the Darfield and Christchurch earthquakes by Tonkin and Taylor through geotechnical investigation, specifically the standard penetration test (SPT). The locations of the measurements correspond to the boreholes from which surface liquefaction observations have been made, as described in section 4.1. When using a method, such as Hazus or Zhu et al. (2015), that predicts a probability of liquefaction, one can interpret the probability as a regional parameter that describes the spatial extent of liquefaction rather than discrete site specific predictions. This is how Zhu et al. (2015) suggest that their model should be interpreted. So for example, one would expect 30% of all sites with a liquefaction probability of 0.3 to exhibit liquefaction and 50% of all sites with a liquefaction probability of 0.5, etc. However, when using liquefaction predictions as means to estimate structural damage over a wide area, it is useful to know not just the number of liquefied sites but also where these sites are. This is particularly important for infrastructure systems since the complexity of these networks means that damage to two identical components can have significantly different impacts on overall

systemic performance depending on the service area of each component and the level of redundancy associated with it.

There are two ways to generate site specific predictions from probabilistic assessments. One approach is to group sites together based on their liquefaction probability, and then randomly assign liquefaction occurrence to sites within the group based on that probability, e.g. by sampling a uniformly distributed random variable. This method is good for ensuring that the spatial extent of the site specific predictions reflects the probabilities and is useful for generating site specific predictions for simulated earthquake scenarios. However, since the locations are selected randomly it has limited value for comparison of predictions to real observations from past earthquakes.

Another method is to set a threshold value for liquefaction occurrence, so all sites with a probability above the threshold are predicted to exhibit liquefaction and all sites with a probability below the threshold are predicted to not exhibit liquefaction. The disadvantage of this approach is that the resulting predictions may not reflect the original probabilities. For example if the designated threshold probability is 0.5 and all sites have a calculated probability greater than this (even if only marginally), then every site will be predicted to liquefy. Conversely if all sites have a probability below 0.5, then none of the sites will be predicted to liquefy. However since there is no random element to the determination of liquefaction occurrence, the predictions are more definitive in spatial terms and hence more useful for the model testing in this study. The threshold approach can also be used to make site-specific predictions based on LPI, by determining a value above which liquefaction is assumed to occur. It is proposed to adopt the threshold approach here, but the value of the threshold needs to be determined. No guidance is given for Hazus, whilst Zhu et al. (2015) propose a threshold of 0.3 to preserve spatial extent, although they also consider thresholds of 0.1 and 0.2. In their original study, Iwasaki et al. (1984) suggest critical values of LPI of 5 and 12 for liquefaction and lateral spreading, respectively. However other localized studies where the LPI method has been applied have found alternative criteria that provide a better fit for observed data as summarized by

Maurer et al. (2014). Since there is uncertainty in the selection of threshold values, this analysis tests a range of values for each model.

The test area is defined as the area for which all input parameters for all models are available. The controlling factor is the V_S profiles, since the kriging technique can only be used to interpolate values between sample points and not extrapolate values outside of the sample extents. Consequently the northern, western and eastern extents of the test area are the same as the extents of the V_S samples. To the south, the controlling factor is PGA, since the data presented in Figure 4.1 exclude the Port Hills and Lyttelton areas. The final test area includes most of Christchurch and a region extending about 10 km to the north of the city towards the town of Kaiapoi, where significant liquefaction was also observed. Surface liquefaction observations for both earthquakes has been obtained from two sources: borehole data provided by consultants Tonkin & Taylor (van Ballegooy et al., 2014), as visualised in Figure 4.3, and maps from the New Zealand Geotechnical Database (New Zealand Earthquake Commission, 2016), which identify the land damage categories (see Table 4.1). The test area is divided into a grid of 100 m x 100 m cells, generating 25,100 test sites and a site is classified as having liquefied if it either contains a borehole observation with land damage category indicating liquefaction or if the NZGD map shows that at least half the cell is occupied by a land damage category indicating liquefaction. However, at some sites no liquefaction observations are available from either source and so these sites are excluded from the analysis. As a result, the test area consists of 20,147 sites for the Darfield earthquake and 22,803 sites for the Christchurch earthquake.

Both the observation and prediction datasets are binary classifications, so standard binary classification measures based on 2 x 2 contingency tables are used to test performance. The contingency table identifies the true positives (TP), true negatives (TN), false positives (FP, Type I error) and false negatives (FN, Type II error). A good predictive model would predict both positive (occurrence of liquefaction) and negative (non-occurrence of liquefaction) results well. Diagnostic scores for each model can be calculated based on different combinations and

functions of the data in the contingency tables. The true positive rate (TPR or sensitivity) is the ratio of true positive predictions to observed positives. The true negative rate (TNR or specificity) is the ratio of true negative predictions to observed negatives. The false positive rate (FPR or fall-out) is the ratio of false positive predictions to true negatives. The best model would have a high TPR and TNR (> 0.5) and low FPR (< 0.5).

The results, presented in a contingency table and associated diagnostic scores, assume a single initial threshold value. Further statistical analysis is undertaken to optimize the thresholds in accordance with the observed data. For a single model, at a specified threshold, the Receiver Operating Characteristic (ROC) is a graphical plot of TPR against FPR. The line representing TPR = FPR is equivalent to random guessing (known as the chance or no-discrimination line). A good model has a ROC above and to the left of the chance line, with perfect classification occurring at (0,1). The diagnostic scores for each model are re-calculated with different thresholds and the resulting ROC values are plotted as a curve for the model. Since better models have points towards the top left of the plot, the area under the ROC curve, AUC, is a generalized measure of model quality that assumes no specific threshold. Since the diagonal of the plot is equivalent to random guessing, AUC = 0.5 suggests a model has no value, while AUC = 1 is a perfect model. For a single point on the ROC curve, Youden's J-statistic is the height between the point and the chance line. The point along the curve which maximizes the J-statistic represents the TPR and FPR values obtained from the optimum threshold for that model.

As well as comparing the performance of simplified models to each other, it is also useful to measure the absolute quality of each model. Simply counting the proportion of correct predictions does not adequately measure model performance, since it does not take into the account the proportion of positive and negative observations, e.g. a negatively biased model will result in a high proportion of correct predictions if the majority of observations are negative. The Matthews correlation coefficient, MCC, is more useful for cases where there is a large difference in the number of positive and negative observations (Matthews, 1975). It is

proportional to the chi-squared statistic for a 2 x 2 contingency table and its interpretation is similar to Pearson's correlation coefficient, so it can be treated as a measure of the goodness-of-fit of a binary classification model (Powers, 2011). From contingency table data, MCC is given by Eq. 4.56.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4.56}$$

Overall the best model will be chosen as the model which results in the highest AUC value that also achieves a reasonable MCC value.

4.4.6.2 Threshold analysis

Model performance is compared by contingency table analysis of true positives (TP), true negatives (TN), false positives (FP, Type I error) and false negatives (FN, Type II error). The true positive rate (TPR) is the ratio of true positive predictions to observed positives. The true negative rate (TNR) is the ratio of true negative predictions to observed negatives. The false positive rate (FPR) is the ratio of false positive predictions to true negatives. An initial set of results is shown in Table 4.13 alongside the corresponding diagnostic scores, using 5 as a triggering threshold for the LPI models; 0.3 as a triggering threshold for the ZHU models; and 0.5 as a triggering threshold for the Hazus. The LPI1, LPI3 and LPIref models are the only models that meet the criteria of having TPR and TNR >0.5 and FPR <0.5, with the LPI1 model performing better despite being based on V_S rather than penetrative ground investigation data (i.e. CPT or SPT results). Table 4.13 shows that all Hazus models are very good at predicting non-occurrence of liquefaction. However, this is only due to the fact that they are predicting no liquefaction all the time, and so their ability to predict the occurrence of liquefaction is extremely poor. The high TNR but relatively low TPR of the three 'ZHU' models indicate that they all show a bias towards the prediction of non-occurrence of liquefaction. The difference between TPR and TNR is indicative of the level of bias in the model and this regard ZHU2, the regional model shows less bias than in ZHU1, the global model, as would be expected. The bias

in the ZHU2 and ZHU3 models is approximately similar although the predictive power of ZHU2 is slightly better.

Table 4.13 – Summary of contingency table data and diagnostic scores for all models using initial threshold estimates

Model	TP	TN	FP	FN	TPR	TNR	FPR
LPI1	6345	25685	9442	1478	0.811	0.731	0.269
LPI2	147	35063	64	7676	0.019	0.998	0.002
LPI3	4287	30578	4549	3536	0.548	0.870	0.130
LPIref	5964	20826	14301	1859	0.762	0.593	0.407
HAZ1	0	35127	0	7823	0.000	1.000	0.000
HAZ2	0	35127	0	7823	0.000	1.000	0.000
HAZ3	0	35127	0	7823	0.000	1.000	0.000
ZHU1	1880	33483	1644	5943	0.240	0.953	0.047
ZHU2	3135	31931	3196	4688	0.401	0.909	0.091
ZHU3	2754	31017	4110	5069	0.352	0.883	0.117

The LPI2 model, using V_{S30} as a proxy, also shows a very strong bias towards predicting non-occurrence. This is expected since V_{S30} generally provides an overestimate of V_S for soil layers at shallow depth. At sites where the soil profile of the top 30 m is characterized by some liquefiable layers at shallow depth with underlying rock or very stiff soil (e.g. in western and central areas close to the inland edge of the sedimentary basin), V_{S30} will be high. Hence, this leads to false classification of shallow layers as non-liquefiable. The LPI3 model with simulated V_S profiles exhibits good performance in the prediction of non-occurrence of liquefaction and correctly predicts just over half of the positive liquefaction observations, indicating bias towards negative predictions. Although the V_S profiles generated through this approach are more realistic than using a constant V_{S30} value, the V_S at each layer is related to V_{S30}. Therefore, at sites characterized by a high V_{S30} value with low V_S values at shallow depths, even using Eq.'s 4.54 and 4.55 may not predict sufficiently low values of V_{S1} to classify the shallow layers as liquefiable. Another factor in the LPI models is the use of the bias-correction factor proposed by

Juang et al. (2005). Whilst this correction factor is appropriate when actual V_S profiles are used, as in LPI1, it may not be appropriate for LPI2 and LPI3 where non-conservative proxies for V_S are used and the resulting misclassification of liquefiable soil layers counteracts the conservativeness of the Andrus and Stokoe (2000) model for estimating CRR. The sensitivity of the models to the correction factor is tested by reproducing the contingency tables for LPI2 and LPI3 with the same threshold values but ignoring the correction factor for the factor of safety, *FS*. These models are referred to as LPI2b and LPI3b and the new contingency table analysis is presented in Table 4.14.

Table 4.14 – Summary of contingency table data and diagnostic scores for LPI models subject to sensitivity test without Juang et al. (2005) correction factors being applied to *FS*

Model	TP	TN	FP	FN	TPR	TNR	FPR
LPI2b	610	34902	225	7213	0.078	0.994	0.006
LPI3b	6068	20509	14618	1755	0.806	0.584	0.416

These results show that not using the bias correction makes little difference to the performance of LPI2, as LPI2b still exhibits an extremely strong bias towards predicting non-occurrence of liquefaction. This is because the purpose of the correction factor is to reduce conservativeness but LPI2 does not need this, since the use of V_{S30} as a proxy for V_S already underestimates liquefaction. For LPI3 however, the difference is more significant. Without the correction factor, the TPR and TNR values for LPI3b reverse, with only just over half negative liquefaction occurrences being correctly forecast. LPI3b therefore exhibits a bias towards positive liquefaction forecasts and so it confers no advantage over LPI3. The results in Table 4.13 and Table 4.14 demonstrate the performance of each model with a single initial threshold value. ROC analysis is used to optimize the thresholds and curves for the eleven simplified models and reference model are generated using the ROCR package in R (Sing et al. 2005), as shown in Figure 4.26. For this study, the threshold for the LPI models is assumed to be a whole number. For the 'HAZ' and 'ZHU' models, the threshold is assumed to be a multiple of 0.05 subject to a minimum value of 0.1, which is the minimum tested by the Zhu et al. (2015).

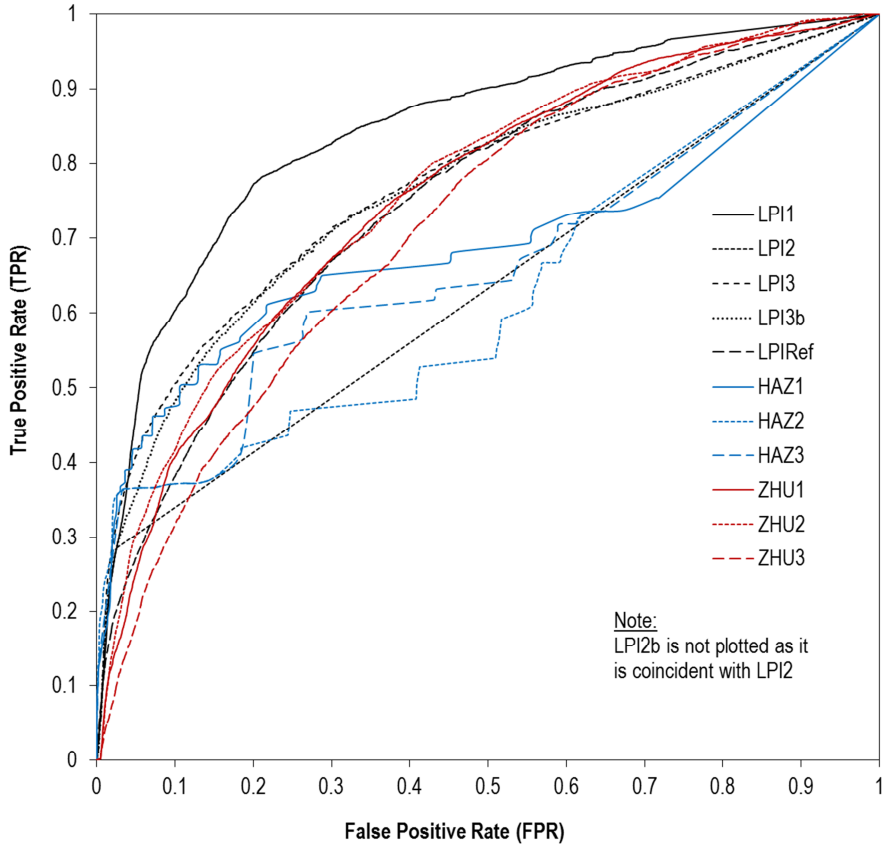

Figure 4.26 – Receiver Operating Characteristic (ROC) curves for prediction models

The AUC values, maximum J-statistics, optimum thresholds and corresponding TPR and TNR values for all models are shown in Table 4.15. With optimized thresholds all the LPI models, except LPI2 and all the ZHU models meet the TPR and TNR criteria (>0.5). All 'HAZ' models and both versions of LPI2 have AUC values closer to the 'no value' criterion, suggesting that the problems with these models lie not just with threshold selection, but more fundamentally with their composition and/or relevance to the case study being tested (noting that the 'HAZ' models have been developed for analysis in the United States). The reason these are to the left of the chance line is because they are predicting non-occurrence of liquefaction at nearly every site and hence they are guaranteed a low FPR value. LPI1 is the best performing model according to both of the ROC diagnostics and although the optimum threshold value of 7 is higher than proposed by Iwasaki et al. (1984), it is within the range for marginal liquefaction – 4 to 8 – proposed by Maurer et al. (2014) and so may be considered plausible.

Table 4.15 – Model quality diagnostics and optimum threshold values for each model from ROC curves

Model	AUC	*J*-statistic	Threshold	TPR	TNR	MCC
LPI1	0.845	0.573	7	0.774	0.799	0.480
LPI2	0.630	0.122	1	0.131	0.991	0.269
LPI2b	0.630	0.206	1	0.224	0.982	0.348
LPI3	0.772	0.420	4	0.581	0.839	0.380
LPI3b	0.766	0.411	10	0.617	0.797	0.357
LPIref	0.748	0.366	6	0.689	0.678	0.290
HAZ1	0.679	0.238	0.1	0.073	0.999	0.238
HAZ2	0.608	0.316	0.1	0.134	0.997	0.316
HAZ3	0.661	0.315	0.1	0.133	0.998	0.315
ZHU1	0.753	0.355	0.1	0.556	0.799	0.311
ZHU2	0.760	0.371	0.1	0.767	0.604	0.287
ZHU3	0.718	0.306	0.1	0.712	0.594	0.237

The two versions of the LPI3 model perform similarly and have reasonable diagnostic scores but LPI3, with the correction factor applied, produces a more plausible optimum threshold value of 4. It is noted however that although the optimum threshold for LPI3b is 10, the TPR and TNR criteria are met with a threshold of 4 but with a lower model performance and greater positive forecast bias (J-statistic = 0.344, TPR = 0.806, TNR = 0.538). The ZHU1 and ZHU2 models perform reasonably with AUC values and J-statistics slightly lower than the LPI3 models, but the optimum thresholds are at the minimum of the tested range, confirming the degree to which these models under-predict liquefaction occurrence. The ZHU2 model also meets the TPR and TNR criteria with a threshold value of 0.2, albeit with a greater prediction bias (J-statistic = 0.370, TPR = 0.555, TNR = 0.815). The ZHU3 model, despite being specific to Christchurch, does not perform as well as ZHU1 or ZHU2. The reason for this anomaly may be because the 'ZHU' models were calibrated to preserve the extent of liquefaction rather than

to make site-specific predictions. The absolute quality of models is tested by calculating MCC. In the preceding analysis, the best performing model is LPI1 and this has a value of MCC = 0.48. The correlation is only moderate, but nevertheless indicates that the model is better than random guessing. As part of a rapid assessment or desktop study for insurance purposes, this may be sufficient.

4.4.6.3 Conclusions

This analysis compares a range of simplified desktop liquefaction assessment methods that may be suitable for use by the insurance sector where seismic risk assessments may be conducted over large areas and where data availability and resources are key constraints. It finds that the liquefaction potential index, when calculated using shear-wave velocity profiles (LPI1) is the best performing model in terms of its ability to correctly predict liquefaction occurrence both positively and negatively. Shear-wave velocity profiles are not always available to analysts however and it is notable therefore that the analysis shows that the next best performing model is the liquefaction potential index calculated with shear-wave velocity profiles simulated from USGS V_{S30} data (LPI3). Since it is based on USGS data, which is publicly accessible online, this method is particularly attractive to those undertaking rapid and/or regional scale desktop assessments.

The Hazus methodology for estimating liquefaction probabilities performs poorly irrespective of triggering threshold. This is significant since Hazus methodologies (not only in respect to liquefaction) are often used as a default model outside of the US when no specific local (or regional) model is available. Models proposed by Zhu et al. (2015) perform well and since they are also based on publicly accessible data, represent another viable option for desktop assessment. The only issue with these models is that they perform optimally with a low threshold probability of 0.1, which may lead to over-prediction of liquefaction when applied to other locations.

4.4.7 Permanent ground deformation

For this case study, only PGDf induced by liquefaction is considered. Landslide-induced PGDf cannot be considered since the geological and topographic data required are not available. For surface fault rupture, it is evident from the source model analysis section 4.3 that only a very small fraction of the overall seismicity is directly under the study area. Surface fault rupture is therefore only likely to affect a very small number of events, and within these events it is only likely to affect a very small number of assets. It is anticipated that surface fault rupture will therefore have negligible impact on the overall results and so is not considered further.

The LPI models selected in section 4.4.6 for predicting liquefaction triggering do not distinguish between settlement and lateral spreading. Therefore it is necessary to predict both horizontal and vertical permanent ground deformations at sites where liquefaction is predicted to occur. The Hazus model (FEMA, 2015) is able to predict both of these parameters and so it is proposed to use this for consistency. A proposed modification to the Hazus model is to use an alternative method for calculating liquefaction probability, inferred from LPI, in place of the function included in the Hazus model.

Since the occurrence of liquefaction at a site is a binary classification variable, it can be modelled by a Bernoulli distribution with probability of liquefaction that depends on the value of LPI. With data from the Darfield and Christchurch earthquakes, functions relating the probability to LPI can be derived using a generalised linear model with probit link function. The probability of liquefaction occurring at site *j*, is given by Eq. 4.57, where Φ is cumulative normal probability distribution function, LPI_j is the value of LPI at site *j*, and *Y** is the probit link function with the form by Eq. 4.58. The link function is a linear model with LPI as a predictor variable and is derived from the individual site observations.

$$P\left[Liq \mid LPI_j\right] = \Phi\left(Y^*\right) \tag{4.57}$$

$$Y^* = \beta_0 + \beta_1 LPI_j \tag{4.58}$$

Figure 4.27 displays the relationships between liquefaction probability and LPI fit by this method for the two best performing LPI models, LPI1 and LPI3, including 95% confidence intervals. The relationships are accompanied by plots of the observed liquefaction occurrence rates, aggregated at each value of LPI.

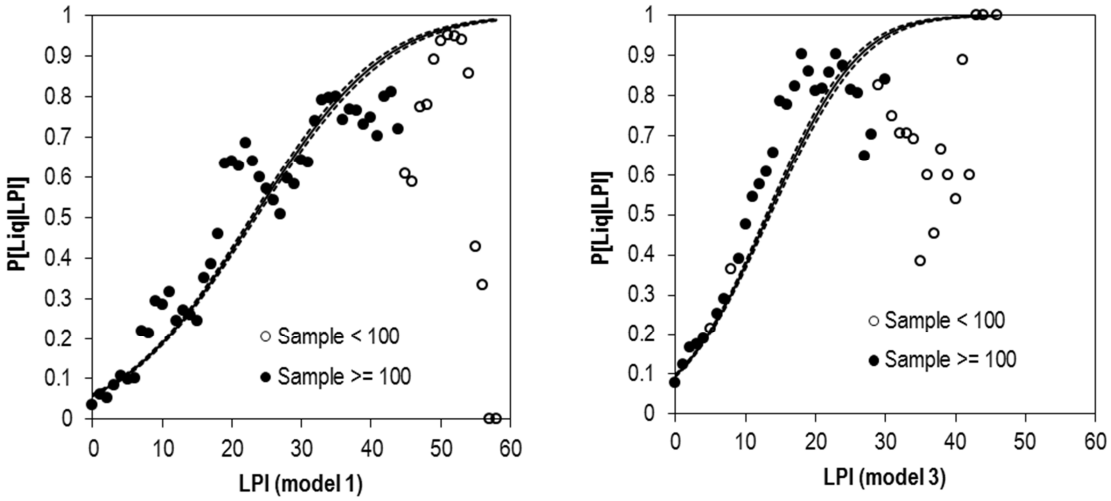

Figure 4.27 – Plots of liquefaction probability against LPI derived from site specific observations by generalized linear model with probit link function for two best performing LPI models. Plots also display the observed liquefaction rates at each LPI value and classified by sample size

The plot for model LPI3b shows greater scatter of observed rates around the fit line than the plot for model LPI1, although in both cases the confidence interval is very narrow, which is a reflection of the large sample size. The confidence interval for LPI1 (±0.0014) is slightly narrower that the confidence interval for LPI3 (±0.0021), indicating that LPI1 is the better predictor of liquefaction probability, just as it is better at predicting liquefaction occurrence by LPI threshold. It is noted in both models though that the observed rates that are furthest away from the fit line are mostly those based on smaller sample sizes (arbitrarily defined here as 100) and which therefore have less influence on the regression – since the use of individual site observations implicitly gives more weight to observations in the region of LPI values for which sample sizes are larger. Furthermore, the observed rates are themselves more unreliable for smaller sample sizes. For example, for model LPI1, observations based on more than 100 samples have an average margin of error of 0.05 whereas the average margin of error for

smaller samples is 0.19. For model LPI3b, observations based on more than 100 samples have an average margin of error of 0.05 and this increases to 0.22 when considering the observations based on smaller samples.

The Hosmer-Lemeshow test (Hosmer and Lemeshow 1980) is a commonly used procedure for assessing the goodness of fit of a generalized linear model when the outcome is a binary classification. However, Paul et al. (2013) show that the test is biased with respect to large sample sizes, with even small departures from the proposed model being classified as significant. They consequently recommend that the test is not used for sample sizes above 25,000. Pseudo-R^2 metrics are also commonly used to test model performance (Smith and McKenna 2013), but these compare the proposed model to a null intercept-only model rather than comparing the model predictions to observations. Although the purpose of the analysis is to develop a function that relates LPI to liquefaction probabilistically, contingency table analysis, with a threshold probability to determine liquefaction occurrence, can be used to test the fit of the model (Steyerberg et al. 2010). Assuming a threshold probability of 0.5, Table 4.16 presents summary statistics from the contingency table analysis of each model and also the coefficients of the corresponding probit link function.

Table 4.16 – Coefficients of link function and summary of contingency table analysis for the two best performing LPI models

Model	β_1	B_0	TPR	TNR	*J*-statistic	AUC
LPI1	0.067	-1.555	0.683	0.869	0.551	0.843
LPI3	0.098	-1.299	0.784	0.856	0.641	0.766

Both models have values of TPR and TNR above 0.5 and the values are of a similar order to those obtained in Table 4.13 and Table 4.14 for the same models but using a value of LPI as a threshold. In particular, values of TNR are now higher, which indicates that the probabilistic LPI model is better at predicting non-occurrence of liquefaction. This is important as over 80% of observations in this study are negative. The difference in values of AUC between Table 4.15

and Table 4.16 are negligible but the J-statistic for LPI3 with a threshold probability of 0.5 is considerably higher than the J-statistic that the optimal threshold found for LPI3 in Table 4.15. This suggests that LPI3 is best implemented as a probabilistic model for liquefaction occurrence. Overall these statistics indicate that both of the probabilistic LPI models proposed are good fits to the observed data, and can therefore be used in the calculation of permanent ground deformations with the Hazus model.

4.5 Step 4 – Fragility functions

This section summarises the procedure and analysis for selecting fragility functions for the key asset types identified in section 4.2. With respect to liquefaction-induced fragility of nodal assets – substations, wells and pumping stations – Hazus is the only existing source for a fragility function and adopts the same function for all three. Therefore the same approach is considered for the analysis of liquefaction-induced fragility of nodal assets here.

4.5.1 Substation fragility

Orion has provided geospatial data for the Transpower GXPs and the district/zone substations. The low levels of damage observed amongst these assets in both earthquakes mean that there are insufficient damage observations to derive new Christchurch-specific fragility functions for substations. In this case, the two options for risk assessment are to either use an existing fragility function as it is or to use a Bayesian updating procedure to modify an existing fragility function to account for the new observed data. The first option is the more straightforward procedure and so is attempted first.

4.5.1.1 Existing fragility functions

Two existing fragility function models for substations that are tested in this study. The first model is the Hazus approach, which treats a substation as a single asset with a single set of applicable fragility functions. The fragility functions have been developed based on empirical data from the United States and expert judgment. For ground shaking damage, there are four

fragility functions representing four damage states: minor, moderate, extensive and complete. The definitions of each of these damage states are shown in Table 4.17. Hazus has different sets of fragility functions for different voltage levels and also depending on whether components within the substation are seismically anchored. The Transpower GXPs and Orion district/zone substations are both most closely matched to the definition of low voltage in Hazus and due to the seismic upgrading that took place in the 1990s (Giovinazzi et al., 2011), it is assumed that all substations are anchored. Hazus fragility functions assume a cumulative lognormal shape and the corresponding median and dispersion values for each damage state due to ground shaking are shown in Table 4.17.

Table 4.17 – Damage state definitions for substations in the Hazus Technical Manual (FEMA, 2015)

Damage state	**Definition**	**Median**	**Dispersion**
Minor	Failure of 5% of disconnect switches or failure of 5% of circuit breakers or building being in minor damage state	0.15 g	0.7
Moderate	Failure of 40% of disconnect switches or failure of 40% of circuit breakers or failure of 40% of current transformers or building being in a moderate damage state	0.29 g	0.55
Extensive	Failure of 70% of disconnect switches, failure of 70% of circuit breakers, failure of 70% of current transformers or failure of 70% of transformers or the building being in an extensive damage state	0.45 g	0.45
Complete	Failure of all disconnect switches, all circuit breakers, all transformers or all current transformers or building being in complete damage state	0.9 g	0.45

For ground failure, Hazus considers the fragility of substations to be the same as for any other building and assumes that buildings can only be undamaged or severely damage due to ground failure. Therefore there is only a single damage state described as ‘at least extensive’, which incorporates both extensive and complete damage. Hazus assumes that the likelihood of minor or moderate damage due to ground failure is small relative to the likelihood of such damage due to ground shaking. Hence the rare occasions when it may occur are tacitly included in the predictions of minor or moderate damage due to ground shaking (FEMA, 2015). The median

value of the fragility function for $PGDf_H$ (lateral spread) is 1.524 m and the median value for $PGDf_V$ (vertical settlement) is 0.254 m. The dispersion for both fragility functions is 1.2.

The second model is that of SYNER-G (Cavalieri et al., 2014c). SYNER-G treats a substation as a collection of smaller inter-connected sub-components, each with their own fragility functions, which have been adapted from work by Vanzi (1996) and based on Italian substations. The fragility functions consist of a single damage state and so represent the probability of failure for that sub-component. The failure probability of the substation is determined using fault tree analysis to identify the potential failure modes of sub-components (i.e. combinations of sub-component failures that can cause complete substation failure) and determining the probability that at least one of these failure modes will occur. Note that the SYNER-G model only adopts this approach for ground shaking. For ground failure, SYNER-G treats substations as a single asset and proposes the use of the Hazus fragility functions. Whilst the SYNER-G approach provides more detailed analysis of substations, it is more complex to apply and requires knowledge of the internal composition of the substation. In this study it is assumed that substations in Christchurch have the same 'typical' layout used in the SYNER-G work. As well as sub-components, this introduces the concept of macro-components. Macro-components are collections of sub-components that exist in series in a substation, such that if one sub-component in the series fails, the entire macro-component fails. A list of sub-components that constitute each macro-component is provided in Table 4.18.

Table 4.18 – Description of constituent sub-components within macro-components in a 'typical' SYNER-G substation

Macro-component	**Sub-components**
Line without transformer	Voltage transformer, coil support, vertical sectionalizing switch, current transformer, circuit breaker, horizontal sectionalizing switch
Autotransformer	2 x horizontal sectionalizing switch, 2 x circuit breaker, 2 x current transformer, 2 x discharger, autotransformer, box
Bars-connecting line	2 x vertical sectionalizing switch, circuit breaker, current transformer, box
Bus	2 x voltage transformer, 2 x bar support

Generic macro-component layouts of transformer and distribution-only substations based on the work of Vanzi (1996) and Cavalieri et al. (2014c) are shown in Figure 4.28. The parameters for the fragility functions are listed in Table 4.19.

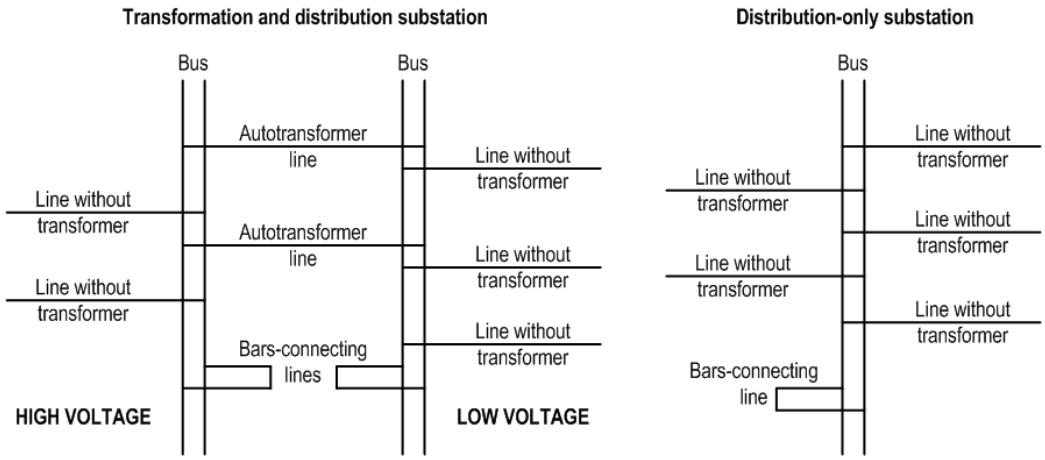

Figure 4.28 – 'Typical' schematic layouts of macro-components in substations for SYNER-G risk assessment (Cavalieri et al., 2014), with flow from left to eight

Table 4.19 – Fragility function parameters for SYNER-G substation sub-components

Sub-component	**Median**	**Dispersion**
Coil support	0.14 g	0.34
Circuit breaker	0.17 g	0.33
Current transformer	0.15 g	0.27
Voltage transformer	0.18 g	0.27
Horizontal sectionalising switch	0.18 g	0.22
Vertical sectionalising switch	0.17 g	0.34
Discharger	0.23 g	0.32
Bar support	0.15 g	0.44
Autotransformer	0.32 g	0.29
Box	0.30 g	0.52
Power supply	0.14 g	0.16

The Orion AMP (Orion, 2009) details the number of transformers and circuit breakers in each substation, and combined with the typical layouts, this information can be used to make inferences on the internal composition of each substation. This has been done by following some assumed rules for the numbers of sub-components and macro-components in each substation, based on the information in Table 4.18 and Figure 4.28, and listed here:

- All lines without transformer have one circuit breaker.
- All autotransformer lines have one autotransformer and two circuit breakers.
- All bars-connecting lines have one circuit breaker.
- The number of bar supports is always four for a transformer substation and two for a distribution-only substation.
- The number of bars-connecting lines is always two for a transformer substation and one for a distribution-only substation.
- The number of incoming lines without transformer is equal to the specified number of transformers in the Orion AMP (Orion, 2009).
- The number of autotransformer lines is equal to the specified number of transformers in the Orion AMP (Orion, 2009).
- The number of outgoing transformer lines is equal to the total number of circuit breakers subtracted by the number of bars-connecting lines and the number of incoming lines without transformer and twice the number of autotransformer lines.

Failure modes for the substations can be considered in terms of individual sub-components or aggregated macro-components. For sub-component analysis, ten failure modes can be identified for a substation (Cavalieri et al., 2014c):

1. Failure of all autotransformers.
2. Failure of all dischargers.
3. Failure of at least one circuit breaker.
4. Failure of at least one horizontal sectionalising switch.

5. Failure of at least one vertical disconnect switch.
6. Failure of at least one bar support.
7. Failure of at least one voltage transformer.
8. Failure of at least one current transformer.
9. Failure of at least one coil.
10. Failure of the power supply.

Failure probabilities for each sub-component can be determined based on ground motion intensity and fault tree analysis can be used to determine the overall failure probability of the substation, by calculating the probability that at least one of the ten failure modes will occur. For macro-component analysis, six failure modes can be identified:

1. Failure of all incoming lines without transformer
2. Failure of all autotransformer lines
3. Failure of all outgoing lines without transformer
4. Failure of at least one bar support
5. Failure of at least one bars-connecting line
6. Failure of power supply

This method requires a two-stage fault tree analysis. The first stage determines the failure probabilities of macro-components based on sub-component failure probabilities, and the second stage determines the substation failure probability from the macro-component failure probabilities.

To determine whether the ground shaking fragility functions are applicable to Christchurch, observed damage is compared to model predictions. Damage to substations was caused by both ground shaking and liquefaction and so descriptive reports by Eidinger and Tang (2012) are used to classify which substations were damaged by ground shaking. Substations damaged by liquefaction only are considered to be undamaged for the purposes of ground shaking fragility assessment. A set of four damage state exceedance probabilities is calculated at each substation

based on observed PGA. A damage state is then assigned to each substation by randomly sampling a uniformly distributed variable *U* between 0 and 1, and then comparing the sampled variable to the set probabilities. If for example a substation has an exceedance probability for 'Minor damage' of 0.6, then it is assigned 'no damage' if the sampled random variable is greater than 0.6. If the exceedance probability for 'Moderate damage' is 0.35, then the structure is assigned to 'Minor damage' if the random variable is sampled between 0.35 and 0.6. The structure therefore has a 0.25 probability that it will be assigned to 'Minor damage'. By repeating this procedure for all substations, a simulated system-wide damage scenario is generated. Since this research focuses on system performance in terms of system connectivity, a damage scenario is simply defined as the number of substations predicted or observed to fail. The SYNER-G method consists of only a 'Failure' damage state but for the Hazus method, a substation is predicted to fail if it is in either the 'Extensive' or 'Complete' damage states (Cavalieri et al., 2014c). This procedure is repeated 1,000 times for each earthquake in a Monte Carlo simulation. The subsequent distributions are then used to determine the expected number of failed substations in each earthquake according to the fragility functions and the likelihood of the two observed damage scenarios occurring given those fragility functions. If, according to predictions, the observed damage scenarios are extreme, i.e. low probability, then one can conclude that the fragility function model is not appropriate for Christchurch. The applicability of the ground failure fragility functions is tested in the same way as the ground shaking fragility functions except that the damage classification is changed, so substations damaged by ground shaking are considered to be undamaged and only substations damaged by liquefaction are considered to be damaged. If the candidate fragility functions are not deemed suitable for Christchurch, the next approach is to use a Bayesian updating procedure with one of the existing fragility functions. The method for this is discussed in section 2.4.3.

4.5.1.2 Existing fragility function testing

The Hazus (FEMA, 2015) and SYNER-G (Cavalieri et al., 2014b) substation fragility models are tested by predicting failure and damage probabilities based on observed PGAs from the

Darfield and Christchurch earthquakes and comparing these to the observed damage. For the Darfield earthquake, only four of the 37 substations were affected by ground shaking, with three substations experiencing damage and one substation experiencing failure. When applying the Hazus model, every substation is estimated to have an exceedance probability for damage of greater than 0.5, whilst for failure, the predicted failure probabilities range from 0 to 0.267, with a median value of 0.056. When applying the SYNER-G model, the results vary depending on whether sub-component or macro-component failure modes are considered. For sub-component failure modes, all but two of the substations have failure probabilities greater than 0.5, whereas for macro-component failure, there is an almost even split with 18 substations having predicted failure probabilities greater than 0.5.

For the Christchurch earthquake only one substation experienced damage and none suffered failure due to ground shaking. Two sites did fail due to liquefaction damage but are excluded from the analysis. Again, when applying the Hazus model, every substation is estimated to have an exceedance probability for damage of greater than 0.5. For failure, the predicted probabilities range from 0 to 0.821, with a median value of 0.245. With the SYNER-G model, sub-component failure mode analysis leads to all but three of the substations have predicted failure probabilities greater than 0.9. Macro-component failure mode analysis is also seen to lead to 26 of the substations having predicted failure probabilities greater than 0.9.

Table 4.20 compares the Hazus model predictions to the observations by pooling together sites within PGA bins of 0.05g intervals. The distribution of damage states within each bin is proportional to the expected damage state distribution modelled by Hazus for the midpoint PGA value of each bin. The bottom two rows compare the totals, and show that generally the Hazus model significantly over-predicts both failure and damage. It is not possible to create an equivalent table for the SYNER-G model, since failure probability depends not just on PGA, but also on the specific layout of each substation. Therefore grouping substations together solely by observed PGA would provide a misleading comparison. Nevertheless, it is clear that the SYNER-G model also over-predicts failure, especially when sub-component failure modes are

applied, generating higher failure probabilities than the macro-component failure modes in every case.

Table 4.20 – Comparison of predictions from Hazus model to observations from Darfield and Christchurch earthquake

Observed PGA (g)	Darfield			Christchurch		
	Failure	Damage only	No damage	Failure	Damage only	No damage
0-0.05	0	0	0	0	0	0
0.06-0.1	0	0	0	0	0	0
0.11-0.15	0	0	0	0	0	1
0.16-0.2	0	4	2	0	0	0
0.21-0.25	2	15	6	1	6	3
0.26-0.3	0	2	1	1	2	1
0.31-0.35	1	2	1	1	2	0
0.36-0.4	0	0	0	1	2	0
0.41-0.45	0	1	0	3	4	0
0.46-0.5	0	0	0	1	1	0
0.51-0.55	0	0	0	1	0	0
0.56-0.6	0	0	0	1	0	0
0.61-0.65	0	0	0	1	0	0
0.66-0.7	0	0	0	2	0	0
Total predicted	**3**	**27**	**7**	**13**	**17**	**5**
Total observed	**1**	**3**	**33**	**0**	**1**	**34**

This phenomenon can be explained with an example. One of the sub-component failure modes is failure of a circuit breaker. If any circuit breaker in the substation fails, then it follows that the entire substation fails. However if macro-component failure modes are considered, then the impact of a failed circuit breaker will depend on the location of the circuit breaker. If the circuit

breaker is in a bars-connecting line, then it will also cause failure of the entire substation. However, if the failed circuit breaker is on an autotransformer line, then it will cause failure of that line, but because there is redundancy amongst autotransformer lines, it will not cause failure of the substation. Consequently the overall likelihood of failure is reduced. Nevertheless, although the application of macro-component failure modes is less conservative, the failure probabilities it generates are still high.

An alternative method for testing model applicability is to randomly assign a damage state to each substation according to its predicted failure probability in order to create a system damage scenario. This is repeated 1,000 times in a Monte Carlo simulation. For the Darfield earthquake, the Hazus model does not generate any damage scenarios with fewer than 15 damaged substations, while the SYNER-G model (with macro-component failure modes) does not generate any damage scenarios with fewer than 12 failed substations. For the Christchurch earthquake, the Hazus model generates a minimum of 21 damaged substations in the Monte Carlo simulation, whilst the SYNER-G model generates a minimum of 26 failed substations. The implication of the simulation results is that with both the Hazus and SYNER-G models, there is a less than 1 in 1000 chance of either of the two observed damage scenarios occurring. This means that neither model can be considered suitable for application to substations in Christchurch. The contrast between the observations and the predictions highlights the very low fragility of the substation. This is attributed to the foresight and accomplishment of the program initiated by Orion to seismically upgrade their substations in the years leading up to the Canterbury earthquake sequence.

4.5.1.3 Bayesian updating

Since neither of the existing models is adequate for Christchurch, it is proposed to apply Bayesian analysis to the Hazus substation fragility function to estimate model parameters for a new Christchurch-specific substation fragility function, using the method described in section 2.4.3. Although the GEM procedure is used to fit generalised linear models, the principles can be applied to any parametric model including cumulative lognormal distributions, which are

more common amongst existing fragility functions. It can be assumed that the natural log of the median for the new function follows a normal distribution whose prior mean is equal to the natural log of the median of the Hazus function. The dispersion parameter has the constraint that it cannot be negative, so it can be assumed to follow a gamma distribution whose mean is equal to the dispersion of the Hazus function. Gamma distributions can be defined by a shape parameter, α_γ, and a rate parameter, β_γ, and the mean of a gamma distribution is the ratio of these two ($\alpha_\gamma/\beta_\gamma$). Approximations of the posterior distributions are obtained using the Gibbs sampling method (Kruschke, 2011), and their mean values are then considered to be the values of the median and dispersion parameters for the new fragility functions, assuming the same cumulative lognormal shape of the original Hazus functions. The Gibbs sampling algorithm is run using the specialised OpenBUGS™ software (Lunn et al., 2009) in conjunction with relevant libraries in R for data input.

Prior knowledge for the new fragility functions are the model parameters for the Hazus substation fragility functions. It is proposed to only produce two fragility curves for Christchurch: one for failure (equivalent to the 'Extensive' or 'Complete' damage states in Hazus) and one for damage only (i.e. the substation remains operational, and is equivalent to the 'Minor' and 'Moderate' damage states in Hazus). Since observed damage data are available for two earthquakes, there are two options to produce and validate a new fragility function: one trained on observations from only the Darfield earthquake and tested on observations from the Christchurch earthquake (the Darfield model); and one trained on observations only from the Christchurch earthquake and tested on observations from the Darfield earthquake (the Christchurch model). Table 4.21 summarises the new parameters for each of these proposed models, estimated in the OpenBUGS software (Lunn et al., 2009). Table 4.22 compares the predictive performance of the new model parameters against observations from the independent dataset. The predictions for each damage state have been made by placing substations into bins based on PGA, as in section 4.5.1.2. Fisher's exact test is used to determine if the differences between the observed and predicted results are significant (Agresti, 2002). This is an alternative

to chi-squared test for cases where some of the predicted values are less than 5 (McDonald, 2014). The null hypothesis in each case is that the difference between observed and predicted data is not significant. Both tests yield p-values greater than 0.05, meaning that the null hypothesis cannot be rejected at the 95% level.

Table 4.21 – New model parameters for fragility functions based on data either only from the Darfield earthquake or only from the Christchurch earthquake

Fragility function	**Failure**		**Damage**	
	Median PGA	**Dispersion**	**Median PGA**	**Dispersion**
Hazus	0.45	0.45	0.15	0.70
Darfield model	0.95	0.33	0.42	0.25
Christchurch model	1.10	0.25	0.64	0.19

Table 4.22 – Comparison of observed damage data to predictions from new models (note: predictions for Darfield earthquake are from Christchurch model and predictions for Christchurch earthquake are from Darfield model)

Damage state	**Darfield earthquake**		**Christchurch earthquake**	
	Observed	**Predicted**	**Observed**	**Predicted**
Failure	1	0	0	0
Damage	3	0	1	7
No damage	33	37	34	28
p-value		0.115		0.055

However, both datasets are relatively small and ideally a new Christchurch-specific fragility function should be based on as much data as possible. The results in Table 4.22 show that the data from both events produce models that can be validated with independent data and so it is proposed that a new fragility function developed using the entire dataset (Darfield and Christchurch observations combined) would also generalise well to independent data. The final parameters for each damage state are shown in Table 4.23, alongside the new model predictions. It is notable that the final parameters are similar to those estimated in Table 4.22 when based

solely on data from the Christchurch earthquake. Therefore, although there are no independent data with which to test the new combined functions, it is assumed that if the Christchurch-only function generalises to independent data, as concluded from Table 4.22, then the combined function must also do so.

Table 4.23 – Model parameters for final substation fragility function and comparison with observed data

Damage state	**Median PGA**	**Dispersion**	**Darfield earthquake**		**Christchurch earthquake**	
			Observed	**Predicted**	**Observed**	**Predicted**
Failure	1.11	0.47	1	0	0	0
Damage	0.61	0.36	3	2	1	7
No damage	n/a	n/a	33	35	34	28
p-value				0.674		0.055

4.5.2 Electrical conduit fragility

Electrical conduits in Christchurch are either overhead lines or buried cables. Vanzi (1996) observed that the transmission towers that support overhead lines are designed for severe wind loading and designed to withstand the asymmetrical pull of cables that can result if an adjacent tower collapses for any reason. This latter condition means that in the event of an earthquake, whilst a tower might be damaged or even collapse, it does not affect the functionality of the attached lines. This view is supported by evidence from the Canterbury earthquakes which shows that there were negligible levels of damage to overhead line infrastructure and even where there was, this had no effect on functionality (Transpower, 2011; Eidinger and Tang, 2012; Kwasinski et al., 2014; Orion, personal communication). Since this study focuses on infrastructure system functionality, it is therefore assumed that overhead lines are not seismically vulnerable.

Buried cables however have been shown to be highly vulnerable to earthquakes. Only Park et al. (2006) specifically consider the vulnerability of conduits, by creating fragility curves based on data from the February 2001 M_W 6.8 Nisqually, Washington earthquake. However, these curves

do not distinguish between overhead lines and buried cables and nor do they consider any physical attributes of the conduits that may impact on their fragility. Furthermore, they only relate fragility to types of ground shaking intensity measures that are not predicted in this case study. Therefore it is necessary to develop new functions empirically with the observations from the Darfield and Christchurch earthquakes.

4.5.2.1 Methodology for empirical function derivation

Orion have provided geospatial damage records for 11 kV buried cables, and by analysing this information in conjunction with geospatial earthquake intensity data, it is possible to develop new functions that can relate metrics of physical damage to cables with IMs. Since buried cables are linear infrastructure, the same procedure as the determination of the seismic vulnerability of pipes can be followed. For pipes, damage is measured in terms of repair rate (number of repairs per unit length), rather than in terms of 'damage states' as is common with buildings or infrastructure nodes. It is therefore proposed to measure cable damage in terms of repair rates and empirically derive repair rate functions, considering a range of IMs and cable attributes. The cable attributes of interest are insulation material, conduction material and decade of construction. It is expected based on the work and views of others (Kwasinski et al., 2014; Orion, personal communication) that insulation material is the uniquely critical factor, but the dataset provides an ideal opportunity to test this supposition.

There are a number of intensity measures that can be used to evaluate ground shaking but it is assumed that maximum horizontal PGV is the most relevant to buried infrastructure since it relates to ground strain (Pineda-Parros and Najafi, 2010). PGV has also been shown in the literature to be well-correlated with damage to pipes (Isoyama et al., 2000; O'Rourke et al., 2001). Whilst in some areas of Christchurch ground shaking was the only observed hazard, in other areas both ground shaking and permanent ground deformation were observed. Kwasinski et al. (2014) conclude that the peak ground velocities observed during the Canterbury earthquakes were not sufficiently large to cause strains in 66 kV cables that would induce failure. Therefore, for this analysis it is assumed that ground deformation is the critical hazard

(O'Rourke et al., 2014), and PGV is only expected to be a factor in areas where liquefaction was not observed.

For permanent ground deformation, five different candidate IMs are proposed. These include $PGDf_V$ for vertical ground deformation and $PGDf_H$ for horizontal ground deformation. In this analysis, $PGDf_V$ is defined as the differential vertical settlement imposed on a cable, which is distinct from the total vertical settlement shown in Figure 4.4 and Figure 4.5. For each 5 m cell in the LiDAR raster map, the difference in total settlement is calculated between the cell itself and each of its eight immediate neighbours in a 3 x 3 cell matrix. The differential settlement is then estimated as the maximum of these eight differences in total settlement. Additionally, three IMs are proposed that combine the effect of horizontal and vertical ground deformation: $PGDf_{MAX}$, which is the maximum of $PGDf_V$ and $PGDf_H$; $PGDf_{VECT}$, which is the vector mean of $PGDf_V$ and $PGDf_H$; and $PGDf_{GEOM}$, which is the geometric mean of $PGDf_V$ and $PGDf_H$. The primary purpose of these three combined IMs is to provide a more detailed analysis in areas where lateral spreading occurred, since lateral spreading can induce both horizontal and vertical movements. It is also of interest to assess whether the combined effect may relate better to cable damage. The specific combinations have not been selected for any known physical relationship. All three methods are however commonly applied to the measurement of ground shaking intensity (ALA 2001; Toprak and Taskin 2006; Akkar and Bommer 2007), which is usually recorded at a station in three orthogonal directions before being reported as a single composite value. The formulae for the combined effect IMs are shown in Eq.'s 4.59 to 4.61.

$$PGDf_{MAX} = \max\left\{PGDf_H, PGDf_V\right\} \tag{4.59}$$

$$PGDf_{VECT} = \sqrt{PGDf_H{}^2 + PGDf_V{}^2} \tag{4.60}$$

$$PGDf_{GEOM} = \sqrt{PGDf_H \times PGDf_V} \tag{4.61}$$

The observed ground deformation dataset used in this analysis is at a high resolution – 5 m for vertical and 56 m for horizontal – and so when assigning IM values to cables, it is observed that

most are exposed to more than one value of $PGDf_V$ and $PGDf_H$. There are different ways that this can be addressed to assume a single value for the entire cable, e.g. maximum, mean, median, mode or another statistical permutation of each value observed along its length. For relatively short cables such an approximation may have little influence on the calculations, but for longer cables (e.g. some cables are in excess of 1 km and so would have over 200 separate $PGDf_V$ observations), there may be significant implications. In particular, exposures to very low and very high values of PGDf may be underestimated due to the averaging process, which in turn could lead to conservative estimates of repair rates at these values. Conversely, exposures to moderate values of PGDf may be overestimated, leading to an underestimation of repair rates. An alternative is to discretise the cables according to the PGV contours and/or PGDf raster cells (e.g. ALA, 2001; Wang and O'Rourke, 2006; Pineda-Porras and Ordaz, 2010; Wang, 2013). This approach allows more precise IM data to be captured in the measurements of exposure, resulting in more reliable repair rates. Discretisation is therefore adopted in this analysis, with cables split into 5 m segments to match the resolution of the most precise IM dataset, $PGDf_V$. Each segment is assigned the PGV value from the closest contour, and the $PGDf_V$ and $PGDf_H$ values from the raster cell in which it is located.

For a particular cable typology and IM combination, the repair rate at each IM level is calculated as the number of observed repairs per kilometre of exposure at that IM level. A repair rate function, in the form of either a linear relationship or a power relationship (see Eq. 2.42), is then generated by performing a linear regression on the series of repair rate (RR) versus IM data points. For PGV, repair rates are only estimated for the Christchurch earthquake since the number of repairs observed in the Darfield earthquake is too small to produce meaningful repair rates. For each zone and PGV combination, the total length of cable exposed and number of repairs is evaluated. The repair rate, RR, for a zone and PGV pair, is the number of repairs, R_{cable}, divided by the length of cable exposed, L_{cable}, as shown in Eq. 4.62.

$$RR\left(PGV \mid Zone\right) = \frac{R_{cable}\left(PGV \mid Zone\right)_{CHRISTCHURCH}}{L_{cable}\left(PGV \mid Zone\right)_{CHRISTCHURCH}} \tag{4.62}$$

However, for liquefaction, the effects are cumulative and therefore the values of $PGDf_V$ and $PGDf_H$ experienced by a cable in the Christchurch earthquake are not independent of the values of $PGDf_V$ and $PGDf_H$ experienced in the earlier Darfield earthquake. Therefore, for PGDf repair rates, each cable should only contribute to the analysis once, which makes the assignment of IM value and liquefaction zone more complex. For cables damaged in the Darfield earthquake, the assigned IM value and liquefaction zone is simply the observation from that event. For other cables, the assigned IM value is the cumulative deformation after the Christchurch earthquake and liquefaction zones are assigned based on a hierarchy. A cable is classified as being in zone D (and by extension zone B) if it is located in an area where lateral spreading was observed in either event. A cable is classified as being in zone C (and zone B) if it is located in area where settlement was observed in either event but no lateral spreading was observed. All other cables are classified as being in zone A. For a given zone and PGDf combination, the repair rate is the number of repairs observed divided by the total cable length exposed.

When deriving repair rate functions for pipes from the Canterbury earthquakes, O'Rourke et al. (2012) use a screening criterion to determine which repair rate data points should be included in the regression, since some may be unreliable due to being based on a small number of faults or small measured area. The principle of the criterion is to calculate the observed repair rate and subsequently determine what is the minimum total cable length required to be statistically confident in the reliability of the repair rate. The latter is defined by the authors as a probability of 0.94 of observing at least two repairs if the distance interval between repairs is assumed to follow a Poisson distribution. The smaller the observed repair rate, the larger the exposure length needs to be. If the exposure on which a repair rate is calculated is below the minimum length, the data point is excluded from the analysis. The formula for the minimum length ($L_{\min}$) is shown in Eq. 4.63.

$$L_{\min} = -\ln(0.01) / RR \tag{4.63}$$

Additional screening criteria are applied to the repair rate observations, including an absolute minimum exposure of 1 km (O'Rourke et al., 2012), in order to limit the influence of potentially unusual localised conditions that only affect small lengths of cable. Furthermore, a minimum of two observed repairs per IM value is included as a condition, since the objective is to calculate a rate. Due to these criteria and the need to ensure that each repair rate observation is based on a sufficient number of faults and total cable length, the regression is performed on the observed repair rates across IM bins rather than for unique IM values. The bin width for PGV is 5 cm/s and the bin width for PGDf is either 0.05 m or 0.1 m, with the width selected in order to maximise the number of data points that meet the screening criteria.

In order to provide a measure of the uncertainty of each repair rate observation, lower and upper bounds of the 95% confidence intervals are calculated by adapting the method of Ulm (1990) and Dobson et al. (1991) for confidence intervals around a Poisson mean, as shown in Eq.'s 4.64 to 4.65. The confidence interval is a function of exposure length, i.e. the greater the exposure, the smaller the confidence interval.

$$RR_{obs\_low} = \frac{\left(\chi^2_{0.975,2\cdot R_{cable}}\right)}{2 \cdot L_{cable}} \tag{4.64}$$

$$RR_{obs\_upp} = \frac{\left(\chi^2_{0.025,2(R_{cable}+1)}\right)}{2 \cdot L_{cable}} \tag{4.65}$$

Another source of uncertainty is in the regression procedure itself. The regression confidence interval at a specified IM value, CI_{IM}, is given by the formula in Eq. 4.66, where RR_{IM} is the mean estimate of the repair rate from the fitted regression model at that IM value, $\overline{\text{IM}}$ is the mean of the observed IM values used in the regression, n_{RR} is the number of observations used in the regression, S_{yx} is the standard error of the repair rate estimates from the regression and t_{crit} is the critical t-statistic with n-2 degrees of freedom (df) and significance level, α_{RR} = 0.05.

$$CI_i = RR_i \pm t(\alpha, df) S_{yx} \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \tag{4.66}$$

Although fragility of linear infrastructure is commonly expressed in terms of repair rates, when system functionality is the primary objective, it is also useful to know the probability that a cable will fail. Since in repair rate analysis the spatial distribution of repairs along a cable is assumed to follow a Poisson distribution (Hwang et al., 1998; Adachi and Ellingwood, 2008), the repair rate is the mean of that distribution for a 1km length of cable. Assuming that cables cannot conduct electricity if they experience one or more faults, then the probability that a cable will fail is equal to the probability that the cable will have at least one repair (Eq. 4.67). Rearranging the probability mass function of the Poisson random variable for zero repairs gives Eq. 4.68, for the probability of failure of a cable of length, L_{cable} km, with repair rate, RR.

$$P[fail] = P[repairs > 0] = 1 - P[repairs = 0] \tag{4.67}$$

$$P[fail] = 1 - e^{-RR \cdot L_{cable}} \tag{4.68}$$

Repair rate is a function of IM, so it is then possible to generate fragility curves for different cable typologies that predict failure probability against IM, similar to those commonly produced for buildings and infrastructure nodes.

4.5.2.2 Repair rate analysis

Repair rate analysis is conducted for the four liquefaction zones summarised in Table 4.2. In each zone, all proposed IMs are analysed. Table 4.24 summarises the observations in each study area for each cable insulation typology. It is important to note that the majority of cables in Christchurch are of the PILCA typology. Consequently, although the final column presents figures for all typologies combined for reference, this is strongly influenced by the PILCA typology. Generally it is observed that across all cable typologies, repair rates are considerably larger in liquefied zones than in zones where ground shaking was the only observed hazard.

Table 4.24 – Observed buried cable repair data by liquefaction zone from both the Christchurch and Darfield earthquakes

Zone		PILCA	XLPE	PILCA HDPE	Other	All
A – No liquefaction	Exposure (km)	2271	639	93	27	3030
	Repairs	64	1	1	1	67
	Repair rate	**0.028**	**0.002**	**0.011**	**0.037**	**0.022**
B – Liquefaction	Exposure (km)	711	121	24	4	860
	Repairs	362	16	10	2	390
	Repair rate	**0.509**	**0.132**	**0.419**	**0.545**	**0.454**
C – Liquefaction, with settlement only	Exposure (km)	649	113	23	3	788
	Repairs	257	14	6	2	279
	Repair rate	**0.396**	**0.124**	**0.266**	**0.586**	**0.354**
D – Liquefaction, with lateral spread	Exposure (km)	62	8	1	0	72
	Repairs	105	2	4	n/a	111
	Repair rate	**1.698**	**0.242**	**2.969**	**n/a**	**1.548**

Furthermore, repair rates in the ground shaking zone are very low. This concurs with the observations of Kwasinski et al. (2014), that the cable materials present in Christchurch should be able to accommodate the ground strains generated by the earthquakes without yielding, as well as the observations from other earthquakes that ground deformation is the primary source of damage to buried cables (Tanaka et al. 2008; Fujisaki et al. 2014). The reliability of the repair rates for the XLPE and PILCA HDPE typologies in the non-liquefaction zone are somewhat uncertain since they are based on a single repair observation. However, they are of a similar magnitude to the repair rates calculated for the PILCA typology and are also similar in the fact that repair rates calculated in the non-liquefaction zone are much smaller than the repair rates calculated in the liquefaction zones. Consequently these repair rates can be considered plausible. Within the liquefaction zone, repair rates are higher in the areas where lateral spreading is observed than in the areas where only settlement is observed, indicating that movement in the

horizontal plane is more damaging to cables. Finally, in general higher repair rates are observed for the PILCA typology than for XLPE or PILCA HDPE typologies. The exception is in the lateral spread zone where the repair rate for PILCA HDPE is very high, although this is based on just a 1km exposure and so may be influenced by unusual local conditions.

Figure 4.29 shows the repair rate observations and fitted regression models for PILCA cables as a function of each of the candidate IMs. For conciseness, only the best performing (as defined by highest coefficient of determination, R^2) of the linear and power relationship models is shown for each IM, although all models are presented on linear axes. Information provided on each plot includes the equation of the best-fit model for predicting a mean value of the repair rate, RR_{MEAN}, R^2, and the p-value for regression significance. Since the repair rate functions are derived from empirical datasets, the observation are characterised by significant natural scatter. Although it is very rare in the literature for empirical functions to be accompanied by estimates of uncertainty (Rossetto et al. 2015), the plots in Figure 4.29 also include information on the regression standard error, SE (in terms of ln RR for the power relationship models) and an error range encompassing one standard error either side of the median prediction of the fitted model. Since the standard error for power relationship models is in terms of ln RR, the standard error becomes a multiplicative factor when converted to natural scale.

The plots show that repair rates do not correlate well with PGV even in the non-liquefaction zone. This supports the observations of Tanaka et al. (2008), Fujisaki et al. (2014) and Kwasinski et al. (2014) that only ground deformation should cause damage to buried cables. Well-correlated and significant regressions are also achieved using $PGDf_V$ and $PGDf_{GEOM}$, which suggests that some cables in this zone may be subjected to subsurface liquefaction. Given that liquefaction is more prevalent when ground shaking is more vigorous and that the zoning study is based on surface evidence of liquefaction only, it is possible that the small number of repairs observed in this zone are the result of zoning misclassification. Due to the small number faults observed, it is not possible to derive a repair rate versus IM model for XLPE, PILCA HDPE or other cable typologies.

Figure 4.30 shows the repair rate observations and fitted regression models for PILCA cables as a function of each of the candidate IMs. In this zone, $PGDf_{GEOM}$ produces the highest R^2 value and the only $R^2 > 0.7$. The regression with this IM is also significant at the 5% level and the error range is not very large relative to the magnitude of the model predictions. Consequently one can conclude that $PGDf_{GEOM}$ is the optimal IM for predicting cable repair rates in liquefied soils.

Figure 4.31 shows the repair rate observations and fitted regression models for PILCA cables as a function of each of the candidate IMs. In this zone, $PGDf_{GEOM}$ produces the highest R^2 value and the only $R^2 > 0.7$. The regression with this IM is also significant at the 5% level. However, since there should be no horizontal movement in an area where only settlement is observed, in practice the $PGDf_{GEOM}$ model can never be applied. This is because the geometric mean of a set of values cannot be calculated if one of the values is zero. The presence of horizontal movements in the empirical dataset is likely to be due to a combination of LiDAR measurement errors or zoning misclassification. As expected PGV and $PGDf_H$ are poor predictors in this zone. Since settlement relates to vertical ground deformation, $PGDf_V$ is the only IM that is physically logical in this zone and would be expected to perform well as predictor. However, although the regression is significant, it performs only moderately in terms of explanatory power, with $R^2 = 0.6$. The error range for the $PGDf_V$ model is relatively narrow however, and so this model may be acceptable. As in zone B, once screening criteria are applied, there are insufficient observations to perform a meaningful regression for other cable typologies.

Figure 4.32 shows the repair rate observations and fitted regression models for PILCA cables as a function of each of the candidate IMs. No IM results in a regression with $R^2 > 0.7$, but the best performing IM is $PGDf_H$ ($R^2 = 0.672$), which is what one would expect in the lateral spreading zone. Also as expected PGV and $PGDf_V$ perform poorly, further indicating the lack of influence that vertical deformation has in areas where lateral spreading is observed. It is notable however that although the R^2 of $PGDf_{GEOM}$ (0.635) is lower than the R^2 for $PGDf_H$, its standard error is also smaller and its 95% confidence interval is narrower indicating lower uncertainty.

Consequently, if one is able to specifically identify areas where lateral spreading will occur, then both $PGDf_H$ and $PGDf_{GEOM}$ could be used as the IM and the final decision rests on the trade-off the analyst is willing to make between explanatory power and uncertainty.

From the preceding analysis it is possible to conclude that ground shaking alone has negligible impact on repair rates compared to liquefaction. Figure 4.29 to Figure 4.32 show that the majority of data points in Zone A have repair rates in the region of 0.01 to 0.1 repairs per km, whilst in Zones B to D, the majority of data points have repair rates greater than 0.1 repairs per km, with many in excess of 1 repair per km. However if an analyst wishes to model risk to buried cables in non-liquefied areas, then PGV is the optimal IM. In areas where liquefaction occurs, $PGDf_{GEOM}$ is the best performing IM, except in areas subjected to lateral spreading, where $PGDf_H$ performs slightly better (but potentially at the cost of increased uncertainty). The repair rate functions associated with the optimal IMs in each zone are summarised in Table 4.25. The uncertainty associated with each model is presented in the corresponding plots in Figure 4.29 to Figure 4.32.

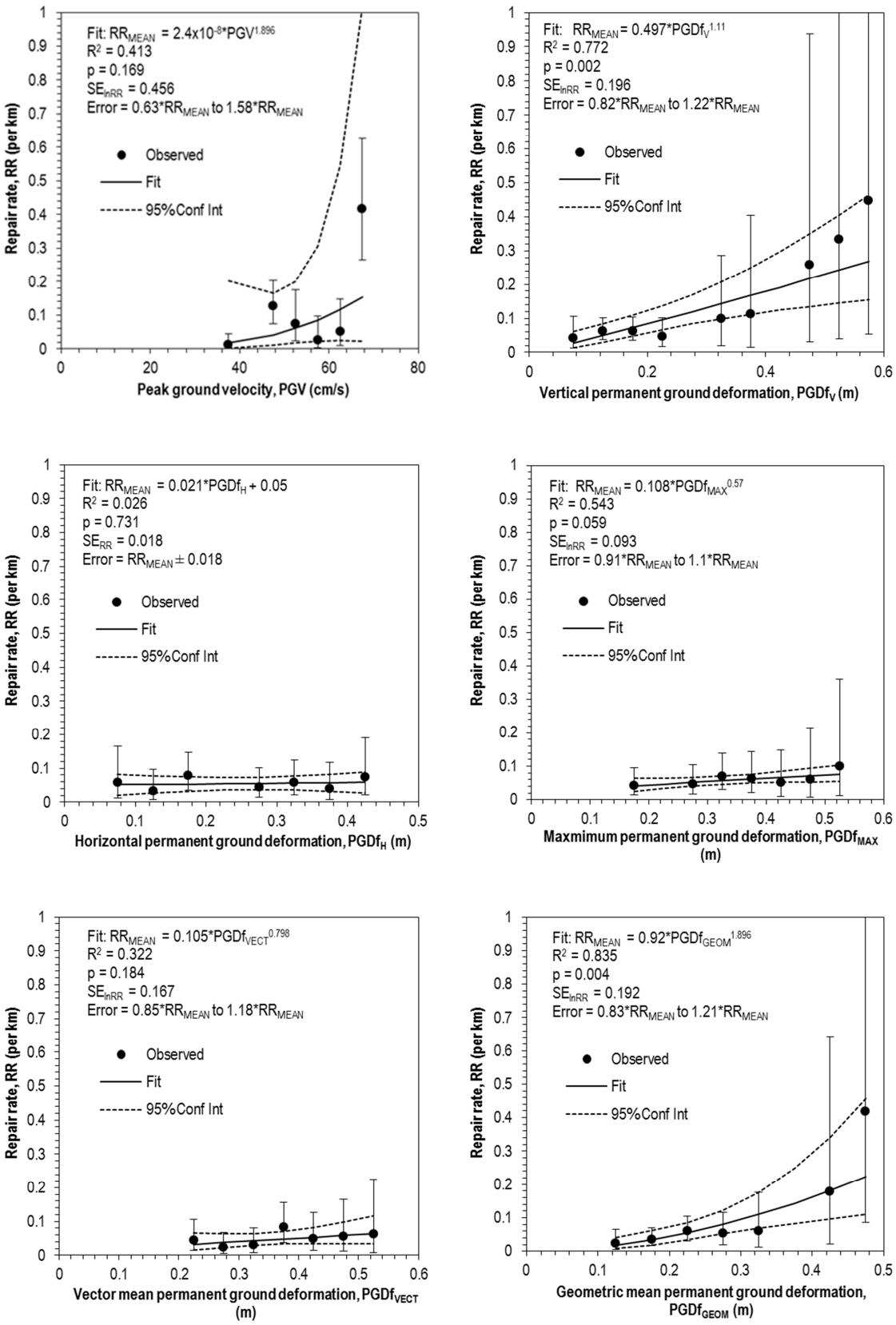

Figure 4.29 – Plots of repair rates versus candidate intensity measures (IMs) for PILCA cables in Zone A (no liquefaction), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

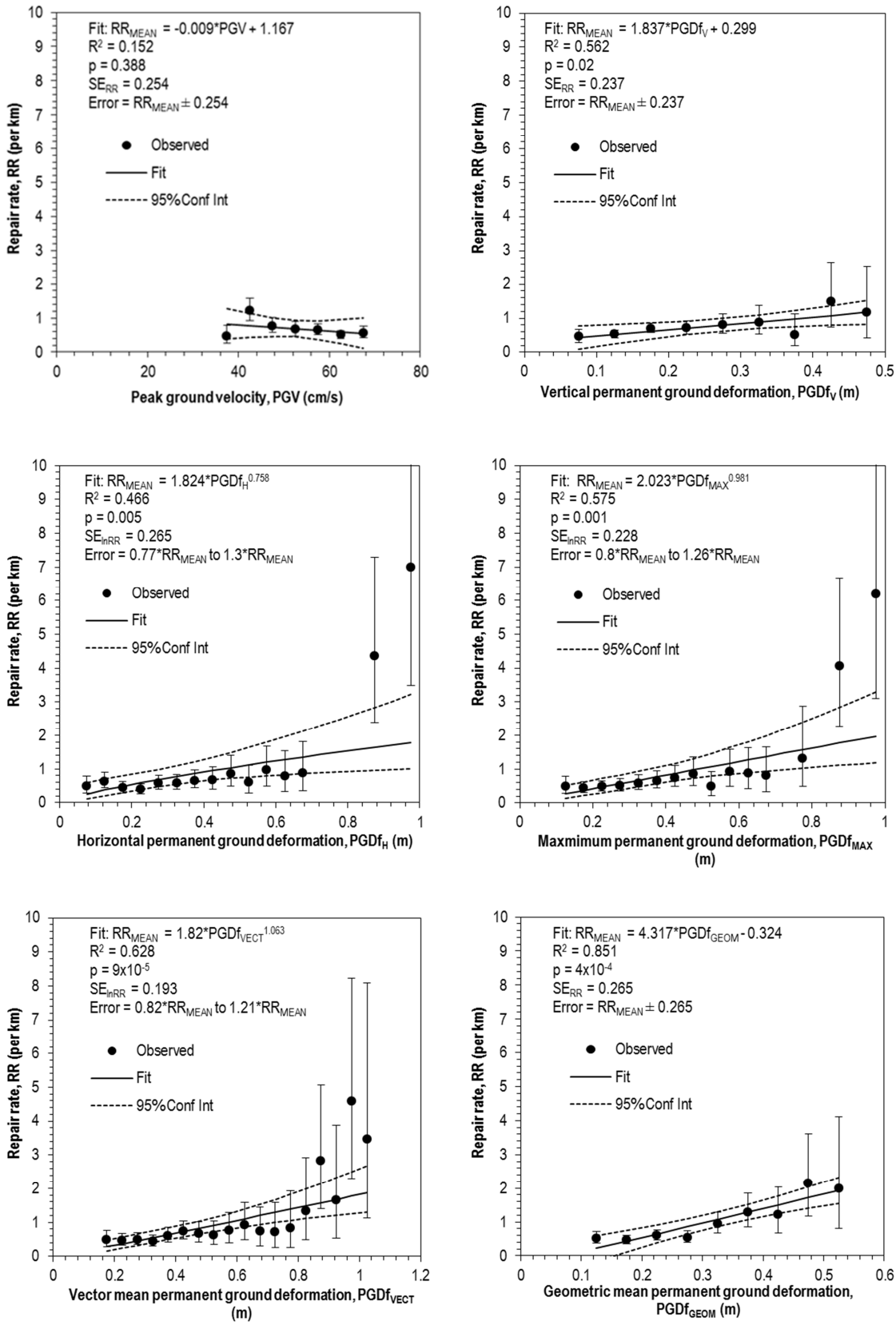

Figure 4.30 – Plots of repair rates versus candidate intensity measures (IMs) for PILCA cables in Zone B (liquefaction), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

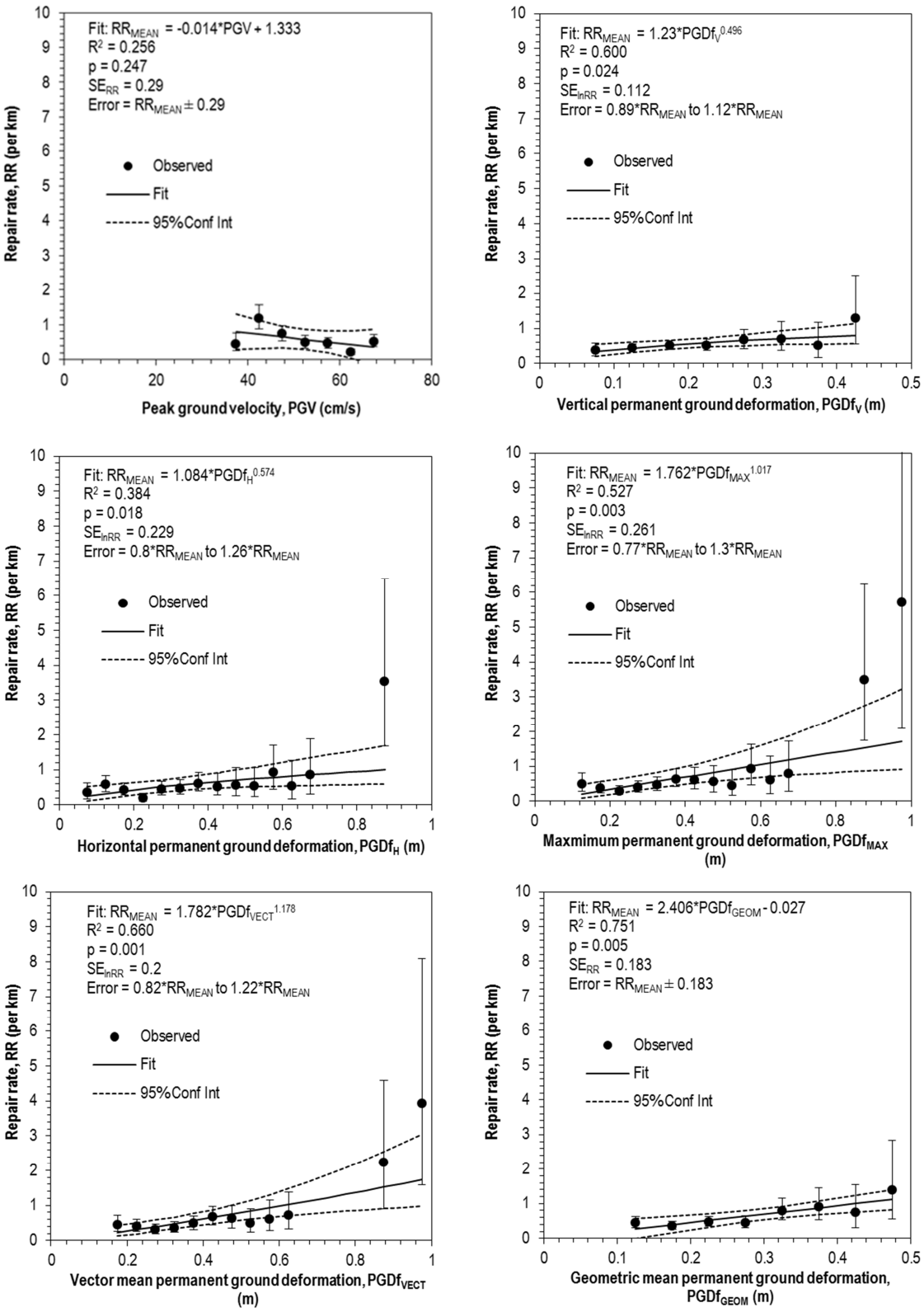

Figure 4.31 – Plots of repair rates versus candidate intensity measures (IMs) for PILCA cables in Zone C (liquefaction with settlement only), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

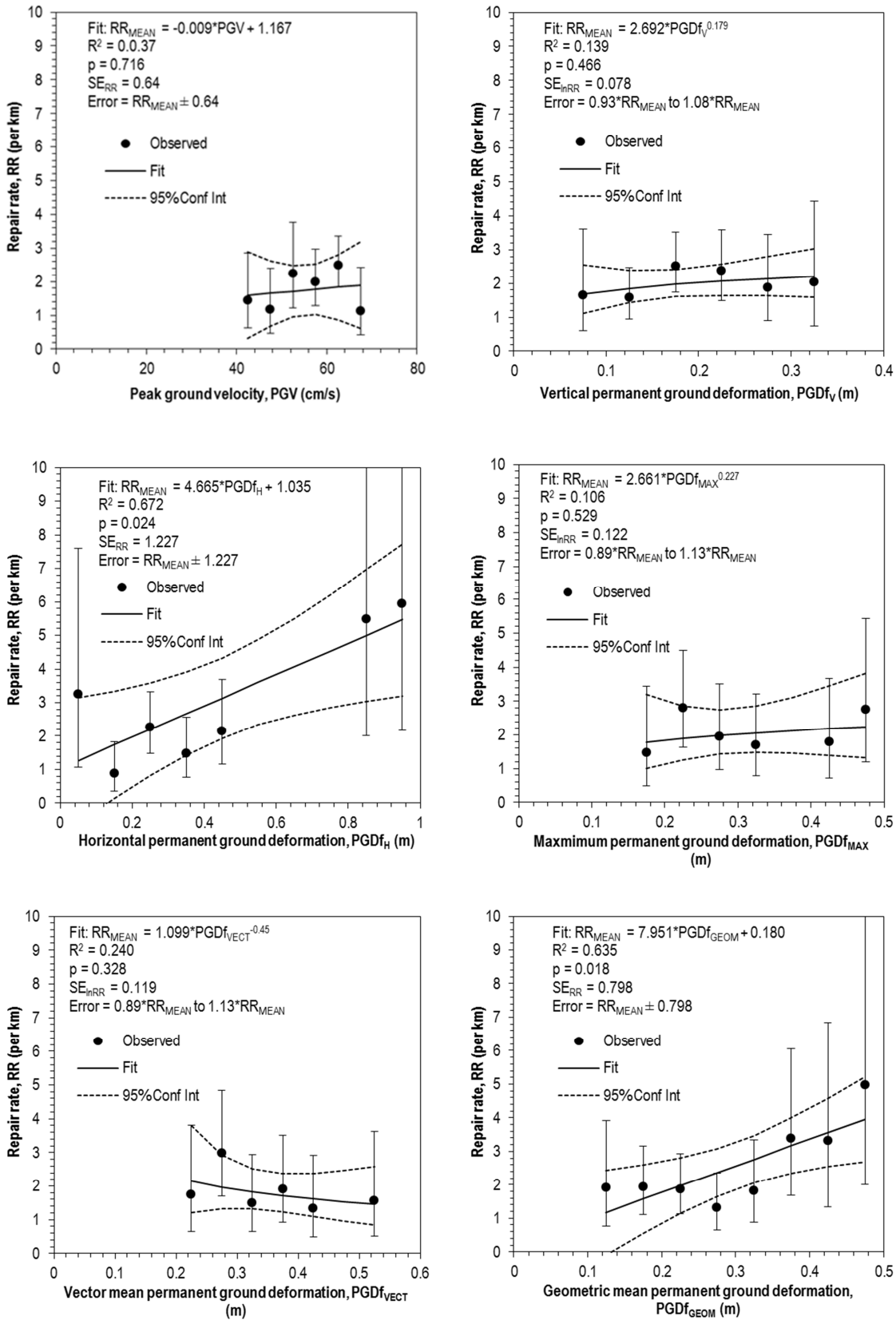

Figure 4.32 – Plots of repair rates versus candidate intensity measures (IMs) for PILCA cables in Zone D (liquefaction with lateral spread), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

Table 4.25 – Optimal IMs and corresponding repair rate functions for each liquefaction analysis zone

Zone	Repair rate function
No liquefaction	No reliable relationship
All Liquefaction	$RR_{MEAN} = 4.317 \cdot PGDf_{GEOM} - 0.324$
Liquefaction – settlement only	$RR_{MEAN} = 1.23 \times PGDf_V^{\,0.496}$
Liquefaction – lateral spreading	$RR_{MEAN} = 4.665 \cdot PGDf_H + 1.035$ or $RR_{MEAN} = 7.951 \cdot PGDf_{GEOM} + 0.18$

The low number of faults in other cable insulation typologies has prevented repair rate versus IM functions being derived for XLPE, PILCA HDPE and 'Other' cable typologies. Yet between them they constitute approximately a quarter of the total cable exposure in Christchurch and must be considered in any risk assessment of the electric power system. As summarized by Kakderi and Argyroudis (2014), it is common with pipe repair rate functions, for the same basic function to be used with a coefficient to account for different material types. A similar approach is proposed for buried cables using the data in Table 4.24.

Taking the PILCA cable repair rate functions as a base model, then the coefficients for alternative typologies can be defined as the ratio of the repair rate in the alternative typology to the repair rate in PILCA cables, not accounting for IM. Therefore to estimate the repair rate for these typologies, one can calculate the repair rate for PILCA cables first and then multiply by the corresponding coefficient. The coefficients for each alternative typology, divided by zone, are shown in Table 4.26.

Table 4.26 – Proposed coefficients for alternative cable typologies to be applied to base PILCA repair rate functions

Zone	XLPE	PILCA HDPE	Other
A	0.06	0.38	1.31
B	0.26	0.82	1.07
C	0.31	0.67	1.48
D	0.14	1.75	0.00
All zones	0.16	0.66	0.68

4.5.2.3 Conduction material

Based on their material properties, Kwasinski et al. (2014) observe that cables that use copper and aluminium as conduction materials (as in the case of Christchurch) should be able to accommodate the moderate liquefaction-induced extensions observed in the two earthquakes, and that other factors primarily affect the fragility of cables. The cable repair dataset includes information on conducting material and so the influence of this factor can be tested. Table 4.27 summarises the repair rates in each zone for cables classified by conducting material, and also for cables classified by their conducting/insulation material combination.

Table 4.27 – Repair rates calculated in each zone for cables classified by conducting material

Zone	Copper All	Aluminium All	Copper PILCA	Aluminium PILCA	Copper XLPE	Aluminium XLPE
A	0.031	0.015	0.032	0.023	0	0.002
B	0.553	0.346	0.553	0.441	0.477	0.120
C	0.417	0.288	0.416	0.366	0.479	0.110
D	1.862	1.104	1.851	1.392	0	0.242
All	0.166	0.078	0.169	0.109	0.052	0.021

Analysing the data for cables classified by conducting material alone, it seems there is a clear difference between copper and aluminium cables, with copper cables approximately twice as vulnerable as aluminium cables in all zones. However, 96% of copper cables are insulated with PILCA, compared to just 61% of aluminium cables. It has been shown in the preceding analysis that PILCA is considerably more vulnerable than other insulation materials and so it is possible that the discrepancy between copper and aluminium as conducting materials is due to the vulnerability of the corresponding insulation rather than due to the influence of the conducting material itself. It is more useful to therefore compare the influence of conducting material between cables with the same insulation material. The dataset for copper XLPE cables is relatively unreliable given that it is based on a low exposure (34 km) and just two repairs. Comparison within PILCA cables is more useful and shows that across all zones, repair rates for copper cables are higher than for aluminium cables. The linear regression procedure for deriving repair rate functions is applied to copper PILCA and aluminium PILCA cables for each zone, using the best performing IMs as determined in the preceding sections. Table 4.28 presents some of the key statistical metrics from the regression analysis.

Table 4.28 – Statistical comparison of repair rate functions derived for copper PILCA and aluminium PILCA cables

Zone	IM	R^2 Copper PILCA	R^2 Aluminium PILCA	t-test p-value
A	PGV	0.626	0.983	0.384
B	$PGDf_{GEOM}$	0.682	0.897	0.559
C	$PGDf_{V}$	0.795	0.163	0.554
C	$PGDf_{GEOM}$	0.729	0.818	0.057
D	$PGDf_{H}$	0.275	0.138	0.146
D	$PGDf_{GEOM}$	0.316	0.666	0.342

R^2 values for the best model fits are presented and indicate that moderate to well-correlated models can be generated for both cable typologies in all zones except zone D, where both IMs result in poor correlations for damage to copper PILCA cables. T-tests are performed to

compare the copper PILCA and aluminium PILCA models in each zone and determine whether they are significantly different. The null hypothesis of each t-test is that there is no significant difference (at the 5% level) between the slopes of regression best fit lines for each typology. In all cases presented in Table 4.28 the p-value is greater than 0.05, so the null hypothesis cannot be rejected. There is insufficient evidence from the data to conclude that conducting material influences repair rates, which corresponds to the observations of Kwasinski et al. (2014). This can be further illustrated by the plots in Figure 4.33. These show that not only do the confidence intervals for the two materials overlap, but that more notably, in each case the best fit line of one material is contained within the confidence bounds of the other. This indicates that there is no significant difference between them.

4.5.2.4 Cable age

A potential factor that may influence the repair rate is the age of the cables. One might expect that older cables would be more vulnerable leading to higher repair rates observed in the data. The dataset provided by Orion includes information on the decade in which each cable was laid. A notable statistic is that 87% of XLPE cables were laid in the 2000s and accordingly, all but one of the faults observed amongst XLPE cable occurred on cables laid during this decade. Consequently a comparison of repair rates with age amongst XLPE cables is not possible. However, this analysis can be performed for PILCA cables and the results are summarized in Table 4.29 and plotted in Figure 4.34, where the age of the cable is taken from the midpoint of each decade to 2010, the year of the first earthquake.

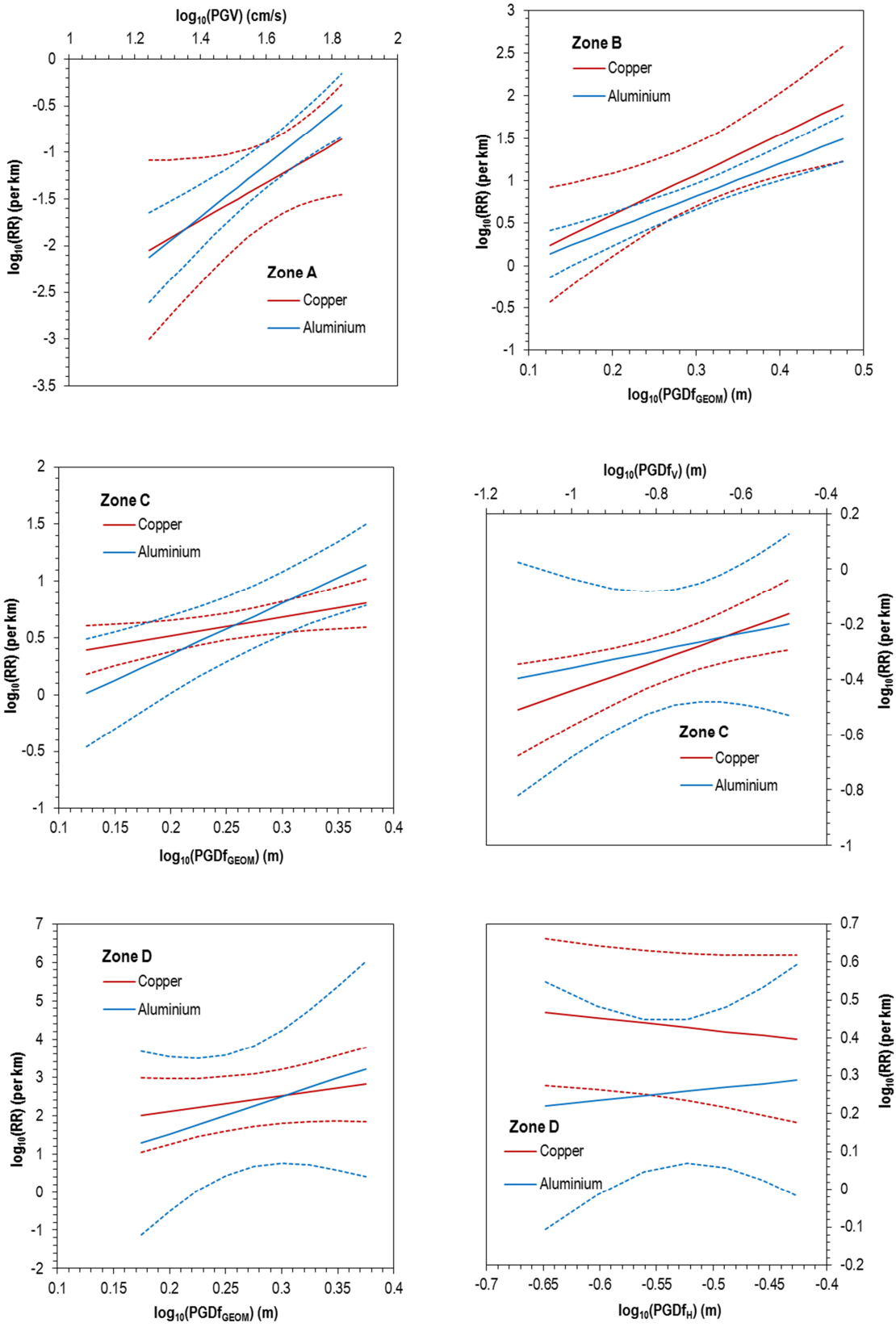

Figure 4.33 – Linear regression model fits and confidence intervals (dashed lines) for copper PILCA and aluminium PILCA cables for selected IMs in each zone

Table 4.29 – Repair rates for PILCA cables in each zone by age (blank cells indicate that the one or more of the screening criteria have not been met)

Decade laid	Age	Zone A	Zone B	Zone C	Zone D	All zones
2000s	5		0.392	0.359		0.084
1990s	15	0.019	0.259	0.159	1.379	0.064
1980s	25	0.043	0.341	0.274	0.991	0.104
1970s	35	0.018	0.641	0.513	2.020	0.181
1960s	45	0.032	0.557	0.406	1.983	0.160
1950s	55	0.037	0.591	0.514	1.466	0.183
1940s	65		0.259	0.221		0.075
1930s	75		0.397	0.169		0.122

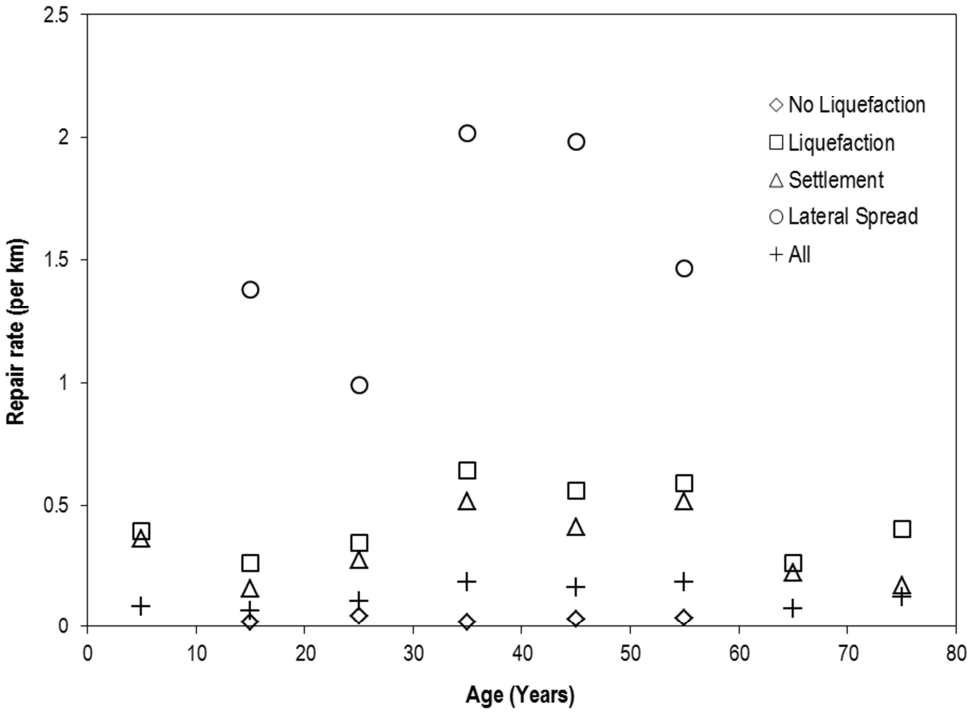

Figure 4.34 – Plot of repair rates versus age in each zone for PILCA cables

The plots do not show any strong trend for repair rate increasing with age in any of the four zones. Linear regression models have been fit for each zone both directly and using logarithmic transformations. None of the models are significant at the 5% level and the highest value of R^2

is 0.213. This indicates that in Christchurch, cable age does not influence its fragility during earthquakes.

4.5.2.5 Fragility curves

Since failure probability depends on length as well as IM, one way to visualise this metric is by plotting a suite of curves on the same axes for different cable lengths. Examples of cable fragility curve suites are shown in **Figure 4.35**, which plots the failure probability of different lengths of cables in Zones C (settlement only) and D (lateral spread), using $PGDf_{GEOM}$ as an IM. For the PILCA curve, repair rates have been calculated using the corresponding functions plotted in Figure 4.31 and Figure 4.32. For the XLPE curve, repair rates have been calculated by applying the relevant material coefficient from Table 4.26 to the PILCA repair rates.

4.5.2.6 Conclusions

This study has used the observations from Christchurch to produce the empirical repair rate functions for buried cables with respect to ground shaking and liquefaction-induced ground deformation. As an empirical dataset, it is characterised by significant natural scatter and this is captured by the inclusion of confidence intervals and uncertainty measurements on the regression plots. Insulation material is a critical factor that influences cable damage as demonstrated by the fact that repair rates in PILCA cables are considerably higher than those observed in XLPE cables. Since there are insufficient damage data to derive specific repair rate functions for materials other than PILCA, all IM analysis has been conducted for PILCA cables and coefficients, derived from the overall repair rates, are proposed to modify the 'base' PILCA functions for other materials. The analysis confirms that liquefaction is the main hazard affecting buried cables, with very low repair rates observed in areas where no liquefaction occurred and even this may be the result of misclassification. There is a poor correlation between repair rate and PGV in this zone, whereas PDf_V and $PGDf_{GEOM}$ show good correlations. This suggests that subsurface liquefaction, which is not accounted for in the liquefaction zoning, may be the primary driver here and that ground shaking alone has only minimal impact on buried cables.

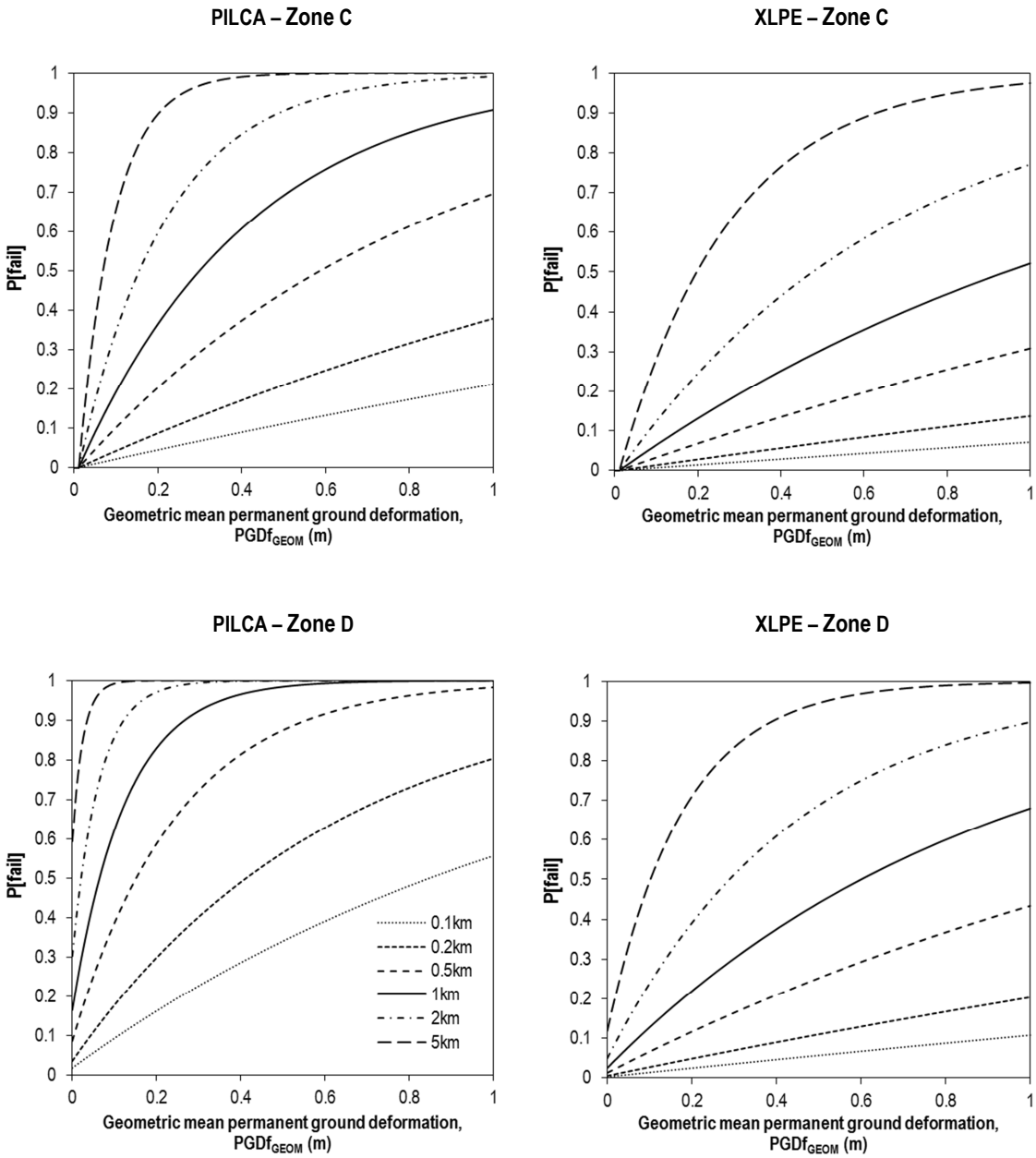

Figure 4.35 – Example suite of fragility curves for PILCA and XLPE cables exposed to settlement only (top row) and to lateral spread (bottom row) measured in terms of PGDf. (Note: legend is the same for all graphs)

Within the liquefaction zone it is notable that lateral spreading is considerably more damaging than vertical settlement alone. In areas where lateral spreading was observed, $PGDf_H$ is the IM that explains the most variance, but $PGDf_{GEOM}$ is only slightly lower in this regard but has smaller uncertainty bounds and so may be considered to be more acceptable to some analysts. In areas where only settlement was observed, $PGDf_{GEOM}$ is the best performing IM in terms of variance explained. However theoretically $PGDf_{GEOM}$ cannot be calculated in a settlement-only

zone and so it is advised to use the function with $PGDf_V$ instead, which has moderate correlations and low uncertainty. When no distinction is made between settlement and lateral spreading, the best performing model is predictably one of the composite IMs, $PGDf_{GEOM}$. Other factors such as conducting material and age have also been considered but there appears to be no trend between repairs and increasing age, while the difference between the repair rates of copper and aluminium cables is not statistically significant. This analysis confirms the findings of Kwasinski et al. (2014) that conduction material should not affect vulnerability.

Although this analysis has been based on data that are characterised by scatter and moderate correlations, they are the first of their kind and there are limited alternatives for addressing buried cables in the literature. Therefore it is proposed to use the optimal functions derived in Table 4.25 for PILCA cables in the seismic risk assessment model and the coefficients in Table 4.26 for other cable typologies. Since the proposed liquefaction triggering models in section 4.4.6.3 do not distinguish between settlement and lateral spreading, it is proposed to use the function in Table 4.25 for zone B (all liquefaction) for liquefaction-induced damage, rather than the individual functions for zones C (settlement) and D (lateral spreading). Also the function for zone B fits the observed data better than the other functions in terms of variance explained (R^2).

4.5.3 Well and pumping station fragility

Damage data for wells and pumping stations are only fully available for the Christchurch earthquake. This dataset identifies which assets failed and which assets were damaged but remained operational, giving two measurable damage states. Since the focus of this study is on system connectivity, establishing fragility functions for complete failure is most critical. In the Christchurch earthquake however, only five wells failed due to ground shaking and these failures occurred across a very narrow range of intensities – 0.19-0.30 g – so developing new fragility functions based solely on this data is of little value.

4.5.3.1 Existing fragility function

Hazus (FEMA, 2015) is the only source for existing fragility functions for wells or pumping stations and includes separate fragility functions for each. In Christchurch many of the wells are attached to pumping stations (known as primary pumping stations). For this type of arrangement, Hazus assumes that the fragility of the primary pumping station is implicit within the fragility function of the associated well. Therefore for fragility analysis, only booster pumping stations need to be considered and not primary pumping stations. Hazus fragility functions have been developed based on empirical data from the United States and expert judgment. For ground shaking damage, there are four fragility functions representing four damage states: minor, moderate, extensive and complete. The definitions of each of these damage states for wells, and the corresponding median and dispersion values are shown in Table 4.30.

Table 4.30 – Damage state definitions for wells in the Hazus Technical Manual (FEMA, 2015)

Damage state	Definition	Median	Dispersion
Minor	Light damage to buildings	0.15 g	0.75
Moderate	Considerable damage to mechanical/electrical equipment or moderate damage to buildings	0.36 g	0.65
Extensive	Building extensively damaged or well pump and vertical shaft being badly distorted and non-functional	0.72 g	0.65
Complete	Building collapsing	1.5 g	0.8

For pumping stations Hazus has different sets of fragility functions for different capacities and also depending on whether the pumping stations have anchored sub-components. The geospatial dataset acquired for booster pumping stations is incomplete in terms of capacity information. Based on the information that does exist, and the fact that all booster pumping stations in Christchurch serve relatively small populations, it is assumed that all stations are ‘small’ by the Hazus definition (<10 million gallons per day). Billings and Charman (2011) state that booster pumping stations performed well in the earthquakes because of earlier seismic upgrades.

Therefore it is assumed that all booster pumping stations have anchored sub-components. The definitions of the damage states for pumping stations, and the corresponding median and dispersion values are shown in Table 4.31.

Table 4.31 – Damage state definitions for booster pumping stations in the Hazus Technical Manual (FEMA, 2015)

Damage state	Definition	Median	Dispersion
Minor	Light damage to buildings	0.15 g	0.7
Moderate	Considerable damage to mechanical/electrical equipment or moderate damage to buildings	0.36 g	0.65
Extensive	Building extensively damaged or the pumps being badly damaged beyond repair	0.66 g	0.65
Complete	Building collapsing	1.5 g	0.8

For ground failure, Hazus considers the fragility of wells and pumping stations to be the same as for buildings and substations. The parameters of the fragility functions are as described in section 4.5.1. The procedure for testing the applicability of these functions to Christchurch is the same as for the procedure used to test the applicability of substation fragility functions, described in section 4.5.1. If the candidate fragility functions are not deemed suitable for Christchurch, the next approach is to use a Bayesian updating procedure with one of the existing fragility functions. The method for this is discussed in section 2.4.3.

4.5.3.2 Existing fragility function testing

The Hazus (FEMA, 2015) fragility functions for wells and pumping stations are tested by predicting failure and damage probabilities based on observed PGAs from the Christchurch earthquake and comparing these to the observed damage. Comparison with observations from the Darfield earthquake is not possible since although the failed assets are known (Eidinger and Tang, 2012), no information is provided with regard to which assets suffered damage but remained operational. Based on data provided by the University of Canterbury, in conjunction with observations reported by Eidinger and Tang (2012), the damage states are known for 124

wells (out of 147) and for all 59 booster pumping stations. Some of these assets have been damaged by liquefaction and so these observations have to be removed from the analysis of the ground shaking fragility function. However, according to the Hazus methodology, 'Minor' or 'Moderate' damage (equivalent here to the 'Damage only' case) caused by liquefaction is intended to be modelled implicitly in the ground shaking fragility function. Therefore only assets that were observed to completely fail due to liquefaction are removed, leaving 113 wells and 57 booster pumping stations. Hazus predictions are generated in the same way as for substations, i.e. by grouping sites into PGA bins, although in this case the bin width is 0.1g. The comparison of predictions with observations is shown in Table 4.32.

Table 4.32 – Comparison of damage predictions from Hazus fragility functions with observations from Christchurch earthquake

Observed PGA (g)	**Wells**			**Booster pumping stations**		
	Failure	**Damage only**	**No damage**	**Failure**	**Damage only**	**No damage**
0-0.1	0	0	0	0	0	0
0.11-0.2	0	1	1	0	0	0
0.21-0.3	1	25	10	0	0	0
0.31-0.4	4	22	4	0	3	0
0.41-0.5	4	14	1	5	13	0
0.51-0.6	2	3	1	7	12	0
0.61-0.7	9	11	0	6	7	0
0.71-0.8	0	0	0	0	0	0
0.81-0.9	0	0	0	3	1	0
Total predicted	**20**	**76**	**17**	**22**	**35**	**0**
Total observed	**5**	**86**	**22**	**2**	**22**	**33**

The results show quite clearly that there is a significant over-prediction of failure and damage by both fragility functions. For wells, a chi-squared test on the totals yields a p-value of 0.0009. For the booster pumping stations, it is necessary to apply a Fisher's exact test since one of the

predicted values is less than 5, and the p-value is less than 1×10^{-14}. This confirms that both existing fragility functions are inadequate for Christchurch.

4.5.3.3 Bayesian updating

Since neither of the existing functions is adequate for Christchurch, it is proposed to apply Bayesian analysis to estimate model parameters for new Christchurch-specific well pumping station fragility functions. The method described in section 2.4.3 is adopted with the Hazus parameters forming the prior knowledge, and observed data from the Christchurch earthquake used. Table 4.33 summarises the new fragility function parameters obtained from this procedure. Since a complete damage dataset is not available for the Darfield earthquake, an independent validation of the model is not possible. However for information, the model predictions for the Christchurch earthquake are shown in Table 4.34.

Table 4.33 – New parameters for well and pumping station fragility functions from Bayesian updating of Hazus functions with observations from Christchurch earthquake

Fragility function		Failure		Damage	
		Median PGA	Dispersion	Median PGA	Dispersion
Well	Hazus	0.72	0.65	0.15	0.75
	Updated	1.39	0.88	0.13	1.29
Pumping station	Hazus	0.66	0.65	0.15	0.70
	Updated	1.22	0.60	0.37	1.45

Table 4.34 – Comparison of existing and updated fragility function predictions with observations from Christchurch earthquake

Model	Wells			Booster pumping stations		
	Failure	Damage only	No damage	Failure	Damage only	No damage
Observed	5	86	22	2	22	33
Hazus	20	76	17	22	35	0
Updated	10	78	25	6	28	23

4.5.4 Pipe fragility

There are numerous existing repair rate functions for pipes but since a large number of repairs were observed across both earthquakes, new Christchurch-specific functions can be developed based on the observed data. The advantage of a locally specific function is that it is able to implicitly capture the effect of pipe attributes that are specific to the locality and are otherwise immeasurable or unobservable, e.g. local construction method or local differences in material production. Although local repair rate functions have been developed previously for Christchurch by O'Rourke et al. (2014), in their analysis each pipe is treated as a single entity with a single observed IM value. In this thesis, it is proposed to discretise the pipes into smaller units and assign observed IM values at a higher resolution. The method for deriving the repair rate functions is identical to the method used for deriving repair rate functions for cables. Based on the data available, the only pipe attribute explicitly analysed is construction material. Fragility functions are also determined for pipes in the same way as for cables but with one difference. With pipes, the presence of a single repair does not automatically mean that the pipe has failed in connectivity terms, since the repair could be the result of a leak, in which case the pipe is still able to convey water. Consequently, whilst a fragility function for pipes would take the same form as Eq. 4.68, it would predict probability of damage rather than probability of failure.

4.5.4.1 Repair rate observations

Repair rate analysis is conducted for the four liquefaction zones summarised in Table 4.2. Table 4.35 summarises the observations in each zone for every pipe material and additionally aggregated repair rates for all brittle materials and all ductile materials. Generally it is observed that across all materials repair rates are considerably larger in liquefied zones than in zones where ground shaking only was observed, and repair rates are also higher where lateral spread occurred compared to where there was only settlement. In contrast to buried cables, the repair rates in the 'No liquefaction' zone are sufficiently high to suggest that ground shaking does

impact pipe vulnerability, which confirms the work of previous studies (Kakderi and Argyroudis, 2014).

Table 4.35 – Observed water supply pipe repair data by liquefaction zone from both the Christchurch and Darfield earthquakes

Material	**Zone A – No liquefaction**			**Zone B - Liquefaction**			**Zone C – Liquefaction, settlement only**			**Zone D – Liquefaction, with lateral spread**		
	Exposure (km)	**Repairs**	**Repair rate**	**Exposure (km)**	**Repairs**	**Repair rate**	**Exposure (km)**	**Repairs**	**Repair rate**	**Exposure (km)**	**Repairs**	**Repair rate**
AC	630	351	**0.557**	263	885	**3.361**	234	692	**2.955**	29	193	**6.623**
CI	118	102	**0.865**	84	232	**2.750**	75	149	**1.977**	9	83	**9.212**
CLDI	5	1	**0.188**	2	2	**1.040**	2	1	**0.577**	<1	1	**5.246**
CLS	23	9	**0.393**	29	58	**2.002**	25	35	**1.417**	4	23	**5.388**
DI	27	9	**0.331**	22	29	**1.327**	20	24	**1.222**	2	5	**2.266**
GALV	119	333	**2.806**	109	841	**7.715**	99	709	**7.173**	10	132	**12.99**
HDPE	541	263	**0.486**	371	419	**1.130**	339	332	**0.979**	32	87	**2.761**
MDPE80	244	45	**0.185**	144	109	**0.756**	123	57	**0.464**	21	52	**2.432**
MPVC	113	4	**0.036**	35	12	**0.345**	31	5	**0.160**	3	7	**2.048**
PVC	179	30	**0.168**	91	81	**0.893**	78	35	**0.450**	13	46	**3.543**
STEEL	21	13	**0.608**	15	38	**2.531**	14	34	**2.478**	1	4	**3.095**
Unknown	23	6	**0.259**	<1	2	**4.845**	<1	2	**9.184**	0	0	**na**
UPVC	66	6	**0.090**	20	9	**0.459**	18	7	**0.388**	2	2	**1.274**
Other	9	65	**7.296**	5	39	**8.336**	4	31	**6.922**	<1	8	**39.98**
Brittle	1224	826	**0.675**	605	2060	**3.423**	536	1597	**2.982**	66	463	**6.985**
Ductile	862	340	**0.395**	583	655	**1.124**	522	483	**0.926**	61	172	**2.826**
All	2118	1237	**0.584**	1189	2756	**2.317**	1062	2113	**1.990**	128	643	**5.038**

Repair rates are particularly high for galvanised iron (GALV) pipes and for ‘Other’ materials, although in the latter case this is based on relatively short exposure. Other brittle materials such

as asbestos cement (AC) and cast iron (CI) also exhibit high repair rates, whilst ductile materials such as ordinary and modified polyvinyl chloride (PVC and MPVC) exhibit the lowest repair rates. The summary at the bottom of Table 4.35 confirms that when aggregated, brittle materials experienced higher repair rates than ductile materials.

4.5.4.2 Repair rate functions

Repair rate functions that predict mean repair rates in terms of IM are developed from the observations by ordinary least squares regression. According to Kakderi and Argyroudis (2014), a set of repair rate functions can be developed either to distinguish between each specific material or to distinguish by material property (brittle or ductile). This analysis develops repair rate functions for each combination of liquefaction zone and material (or material property). Given the twelve identifiable construction materials and the two properties (brittle and ductile) with the four zones, there are 56 potential repair rate functions. For each material/zone combination, the same six IMs that were used for the cable repair rate analysis are tested: PGV, $PGDf_V$, $PGDf_H$, $PGDf_{MAX}$, $PGDf_{VECT}$, $PGDf_{GEOM}$. Since every combination is tested with both a linear and power law relationship as a functional form, there are potentially a total of 672 regressions in this analysis. Many of these do not pass the screening criteria (O'Rourke et al., 2012) described in section 4.5.2.1 for buried cables and therefore do not progress to the stage of the regression being performed. However there are still a total of 328 regressions performed. For conciseness, the full set of results are supplied in Appendix D, whilst this section focuses on the derived repair rate functions that are deemed significant and therefore applicable to the final model, which is defined here as a function with $R^2 > 0.7$ and p-value < 0.05. In some cases, it is possible to derive more than one significant repair rate function for a material/zone combination. In this case the optimum repair rate function for that combination is the one with the highest R^2. Table 4.36 shows that only 15 of the material/zone combinations yield significant repair rate functions. The optimum functions and corresponding observations are plotted in Figure 4.36 to Figure 4.39.

Table 4.36 – Significant pipe repair rate functions derived from Canterbury earthquake sequence observations

Zone	Material	Optimum function			
		Form	IM	R^2	p-value
A	AC	Power	PGV	0.713	0.004
A	HDPE	Power	PGV	0.895	0.0001
A	MDPE80	Linear	PGV	0.813	0.036
A	Brittle	Power	PGV	0.746	0.001
A	Ductile	Power	PGV	0.900	$1x10^{-5}$
B	AC	Power	$PGDf_{GEOM}$	0.807	0.0002
B	CI	Linear	$PGDf_{GEOM}$	0.897	0.001
B	CLS	Power	$PGDf_{H}$	0.777	0.048
B	MDPE80	Power	$PGDf_{VECT}$	0.884	$2x10^{-5}$
B	PVC	Linear	$PGDf_{H}$	0.771	0.021
B	Ductile	Linear	$PGDf_{GEOM}$	0.764	0.002
C	AC	Power	$PGDf_{GEOM}$	0.75	0.001
C	CI	Linear	$PGDf_{V}$	0.841	0.028
D	AC	Power	$PGDf_{H}$	0.877	0.0002
D	MDPE80	Power	$PGDf_{GEOM}$	0.844	0.027

The liquefaction triggering model proposed in this case study (section 4.4.6.3) predicts the occurrence of liquefaction but does not distinguish between liquefaction with settlement only or liquefaction with settlement and lateral spread. Therefore only repair rate functions for zones A and B can be applied in the seismic risk assessment. Only asbestos cement (AC), medium density polyethylene (MDPE80) and the ductile material grouping yield significant repair rates in both zones. In the simplified analysis network proposed for the seismic risk assessment (see Table 4.6), AC is by far the most common material, making up 40% (109 km) of the total network by length, so it is beneficial to have functions specific to this material.

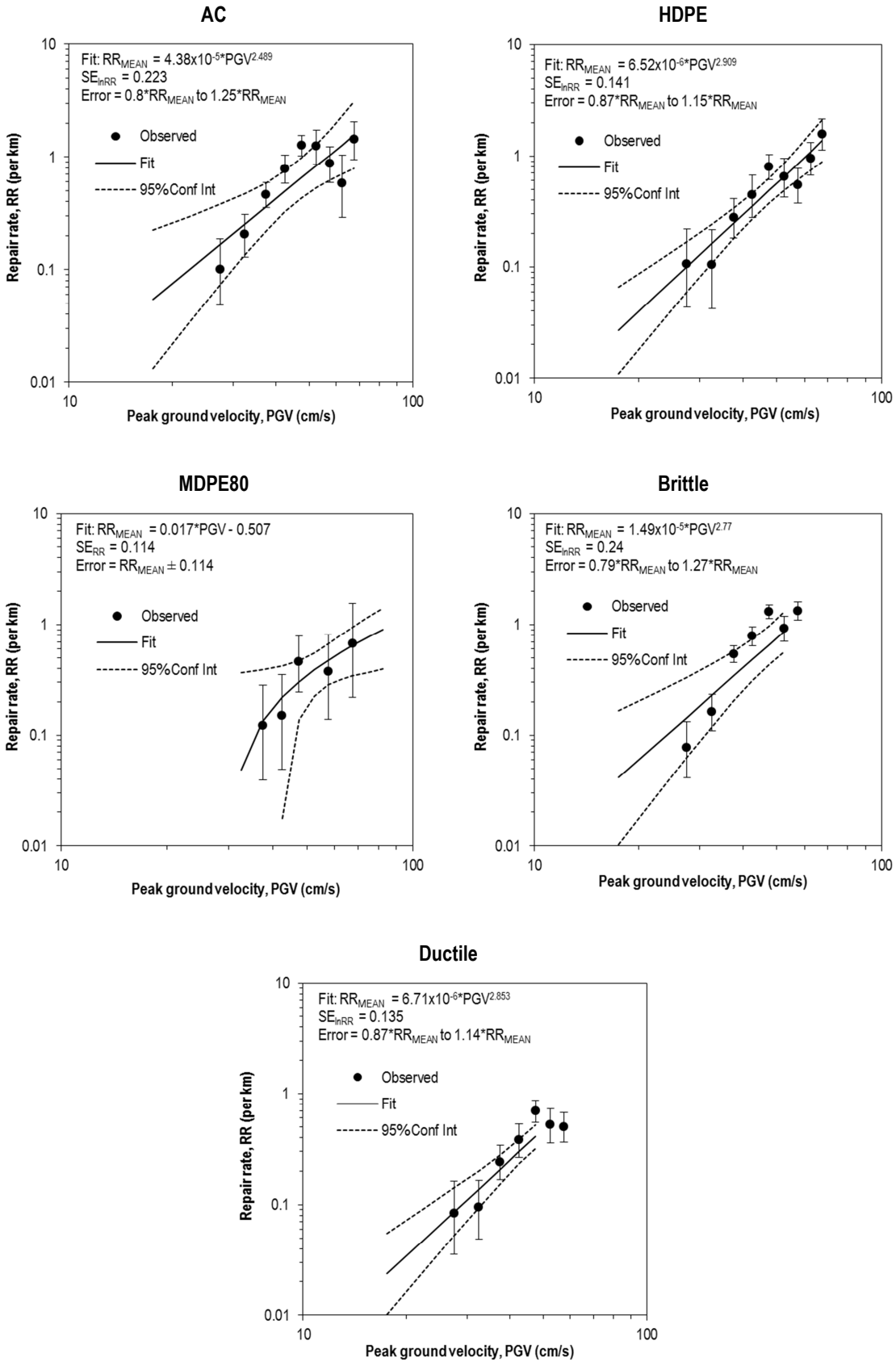

Figure 4.36 – Plots of significant repair rate functions for pipes in Zone A (no liquefaction), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

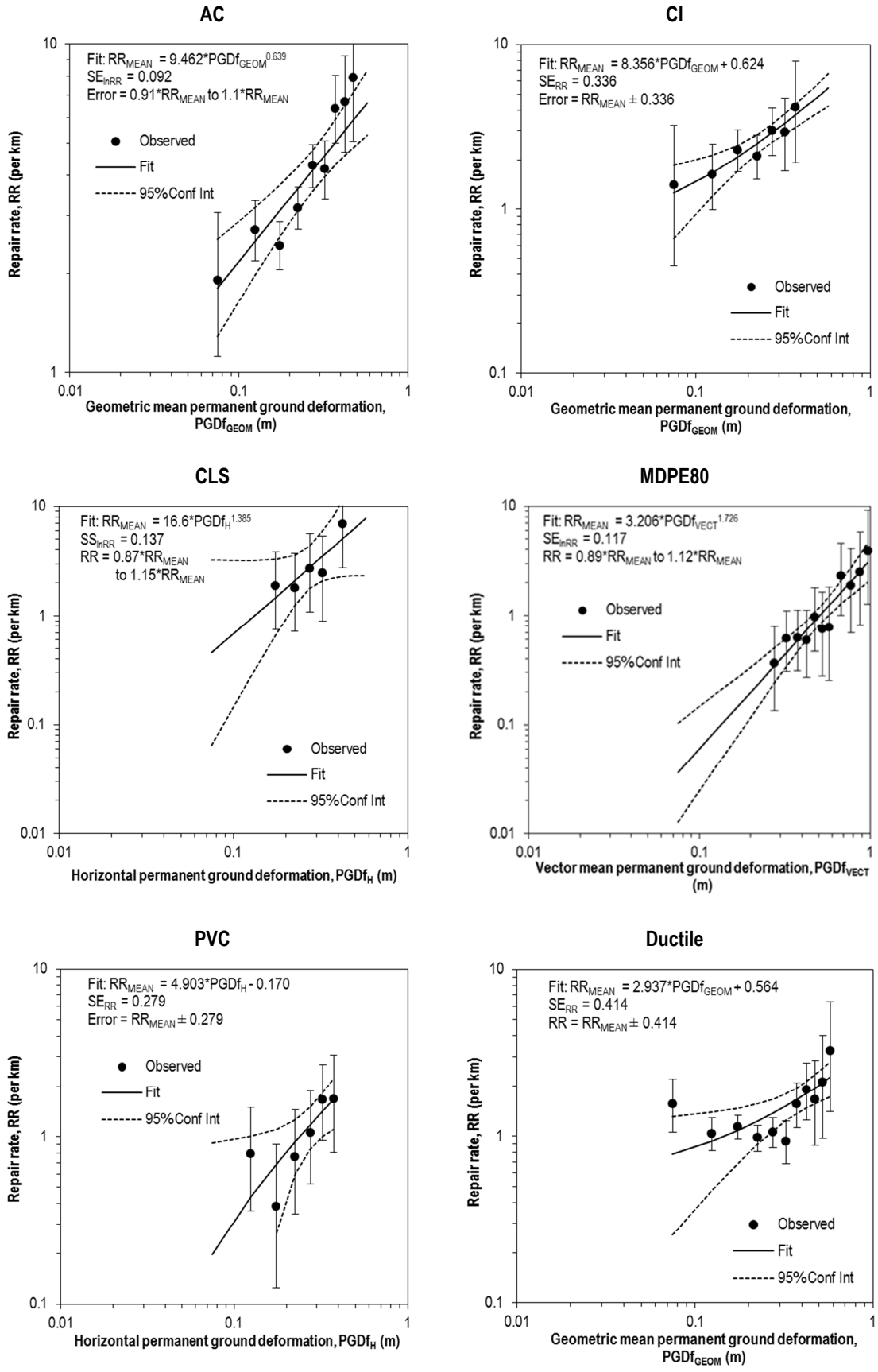

Figure 4.37 – Plots of significant repair rate functions for pipes in Zone B (liquefaction), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

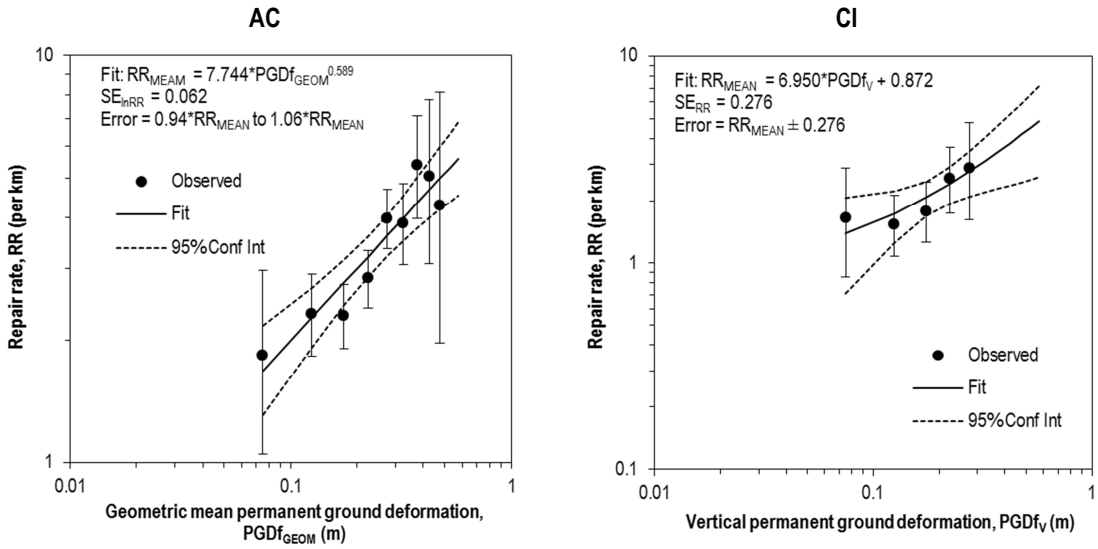

Figure 4.38 – Plots of significant repair rate functions for pipes in Zone C (liquefaction with settlement only), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

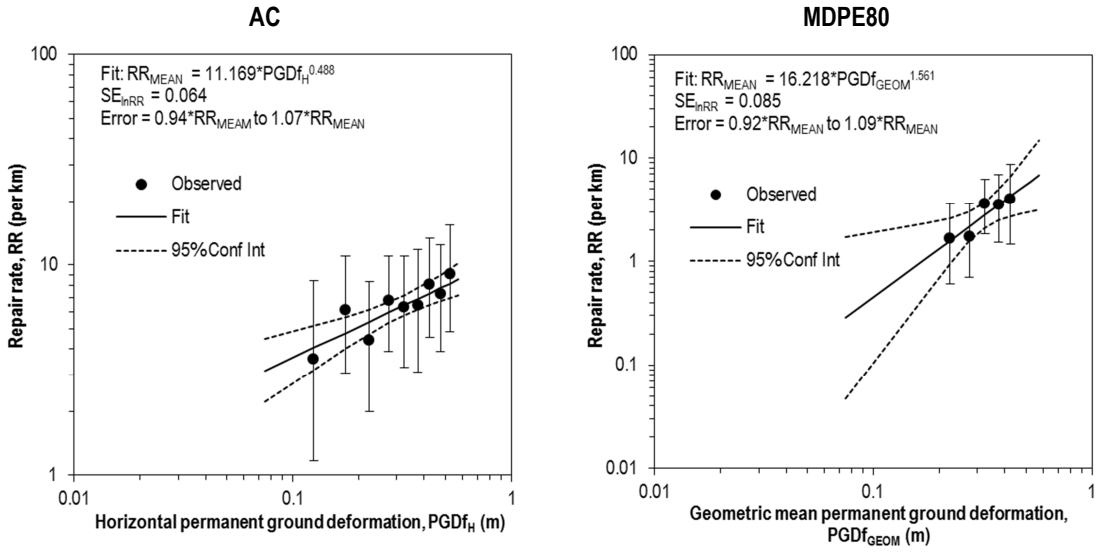

Figure 4.39 – Plots of significant repair rate functions for pipes in Zone D (liquefaction with lateral spread), including Poisson confidence interval around each observation (error bars), best fit linear regression model (solid line) and confidence interval around best fit (dashed line)

In contrast, there is less than 1 km of MDPE80 pipes in the simplified network and so this is of limited value for this particular case study, although may be useful in future studies that take account of the lower hierarchical levels in the Christchurch water supply system. There are significant repair rate functions for high density polyethylene (HDPE) and the brittle material grouping in Zone A and for cast iron (CI), cement-lined steel (CLS) and polyvinyl chloride (PVC) in Zone B. Given the quality of the derived repair rate functions for the most common

pipe material, (i.e. asbestos cement), it is proposed in this case study to apply repair rate functions by material rather than by material property. For other materials that do not yield significant repair rate functions, it is proposed to adopt the same approach as for buried cables. By identifying a base model and applying coefficients to the base model, specific functions for each material are created, using the data in Table 4.35. Taking the AC pipe repair rate functions as a base model, then the coefficients for alternative materials can be defined as the ratio of the repair rate in the alternative material to the repair rate in AC pipes, not accounting for IM. Therefore to estimate the repair rate for these materials, one can calculate the repair rate for AC pipes first and then multiply by the corresponding coefficient. The coefficients for each alternative material, divided by zone, are shown in Table 4.37. For reference, coefficients for the brittle and ductile material properties are also included in Table 4.37.

Table 4.37 – Proposed coefficients for alternative pipe materials to be applied to base AC repair rate functions

Material	Zone A	Zone B	Zone C	Zone D
CI	1.55	0.82	0.67	1.39
CLDI	0.34	0.31	0.20	0.79
CLS	0.71	0.60	0.48	0.81
DI	0.59	0.39	0.41	0.34
GALV	5.03	2.30	2.43	1.96
HDPE	0.87	0.34	0.33	0.42
MDPE80	0.33	0.23	0.16	0.37
MPVC	0.06	0.10	0.05	0.31
PVC	0.30	0.27	0.15	0.53
STEEL	1.09	0.75	0.84	0.47
Unknown	0.47	1.44	3.11	0.00
UPVC	0.16	0.14	0.13	0.19
Other	13.09	2.48	2.34	6.04
Brittle	1.21	1.02	1.05	1.01
Ductile	0.71	0.33	0.43	0.31

4.5.4.3 Fragility curves

Since the probability of a pipe being damaged depends on length as well as IM, one way to visualise this metric is by plotting a suite of curves on the same axes for different pipe lengths. Examples of pipe fragility curve suites are shown in Figure 4.40, which plots the failure probability of different lengths of AC pipe in each of the four zones, assuming the optimum repair rate functions defined in Figure 4.36 to Figure 4.39.

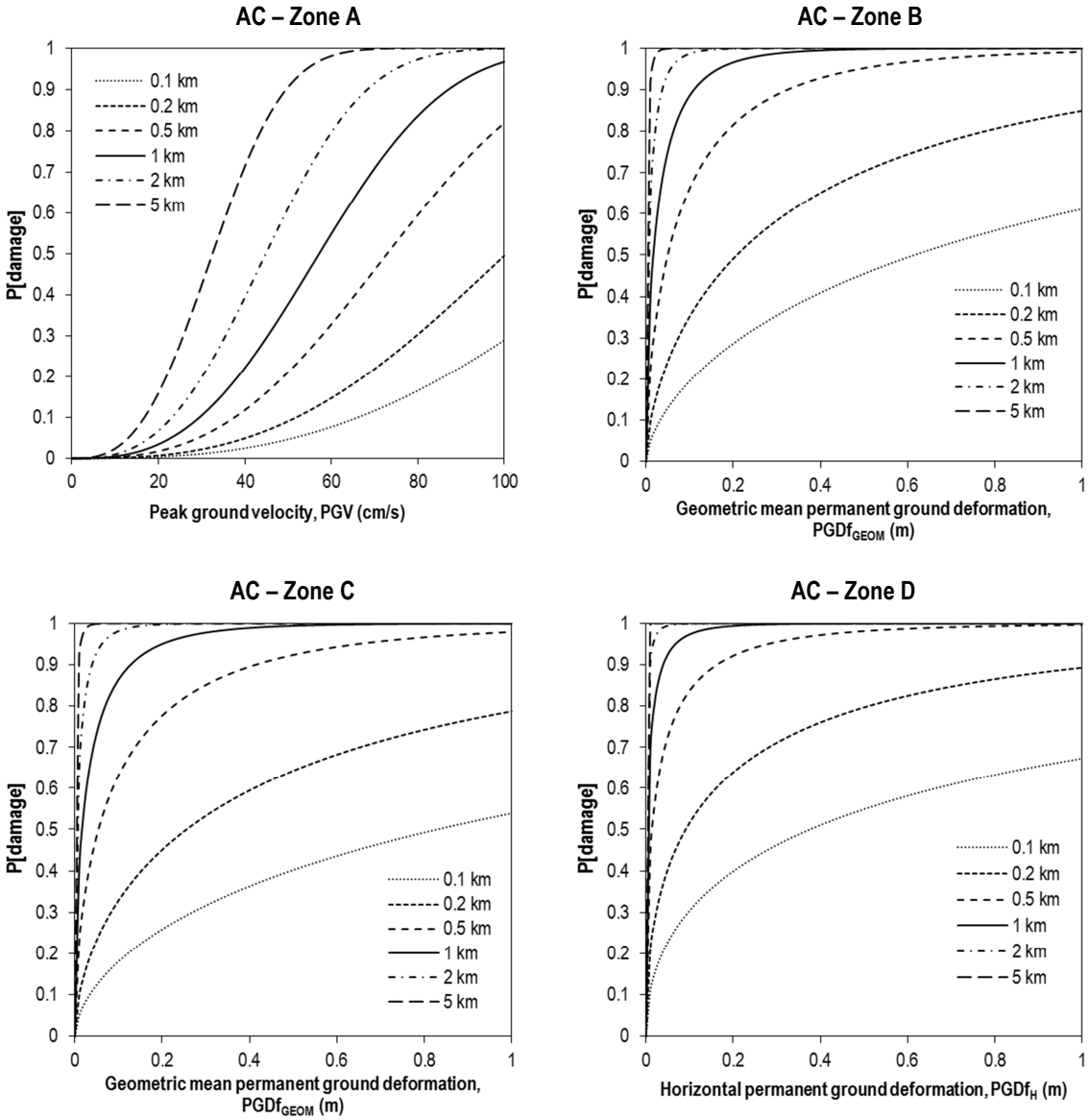

Figure 4.40 – Example suite of fragility curves for AC pipes exposed to different liquefaction zones

4.5.5 Node vulnerability to liquefaction

4.5.5.1 Existing fragility functions

The Hazus (FEMA, 2015) fragility functions for liquefaction-induced failure to nodes are tested by predicting failure probabilities based on observed values of $PGDf_V$ and $PGDf_H$ and comparing these to the observed damage states. It is noted that the Hazus fragility function only accounts for 'Extensive' and 'Complete' damage states (equivalent to failure in this study), since liquefaction-induced damage is assumed to be implicit in the relevant ground shaking fragility function for each asset. When multiple earthquakes occur in sequence, measurements of ground shaking are independent from one earthquake to the next, but measurements of ground deformation are cumulative and hence the observations from each event cannot be treated separately. For each ground deformation metric, there is a single dataset in which an asset appears only once. If an asset failed due to liquefaction in the Darfield earthquake, then the corresponding observed ground deformation metric for that asset is the observation from the Darfield earthquake. Otherwise the observed ground deformation metric is the cumulative deformation across both the Darfield and Christchurch earthquakes. Only assets exposed to liquefaction in either of the two earthquakes are included in the dataset. This is determined from the observed land damage categories reported by van Ballegooy et al. (2014, see Table 4.2). The cause of asset failures (i.e. liquefaction or ground shaking) is determined from reporting by Eidinger and Tang (2012).

For vertical settlement ($PGDf_V$), there are 61 asset observations (45 wells and 16 substations) but for lateral spread ($PGDf_H$) there are only observations from five wells. The lack of observations for pumping stations is as expected since these predominantly serve communities in the Port Hills area, which is built on rocky ground. Table 4.38 shows the predictions from the Hazus fragility functions based on PGDf bin widths of 0.1 m. The results show that the Hazus functions over-predict failure due to vertical settlement and under-predict failure due to lateral spread. A chi-squared test applied to the totals for vertical settlement yields a p-value less than

$1x10^{-4}$, whilst a Fisher's exact test applied to the totals for lateral spread yields a p-value of 0.167, although this needs to be interpreted with caution given the very small sample size.

Table 4.38 – Comparison of failure predictions from Hazus fragility functions with cumulative observations from Darfield and Christchurch earthquakes

Observed PGDf (m)	Hazus Predictions			
	Vertical settlement		Lateral spread	
	Failure	No failure	Failure	No failure
0-0.1	3	27	0	0
0.1-0.2	5	10	0	1
0.2-0.3	3	3	0	2
0.3-0.4	4	2	0	0
0.4-0.5	3	1	0	0
0.5-0.6	0	0	0	1
0.6-0.7	0	0	0	0
0.7-0.8	0	0	0	0
0.8-0.9	0	0	0	0
0.9-1.0	0	0	0	1
Total predicted	**18**	**43**	**0**	**5**
Total observed	**5**	**56**	**3**	**2**

4.5.5.2 Bayesian updating

Since the existing function for vertical settlement is inadequate for Christchurch, it is proposed to apply Bayesian analysis to estimate model parameters for new Christchurch-specific function. The method described in section 2.4.3 is used with the Hazus parameters taken as the prior knowledge and observed cumulative data from the Darfield and Christchurch earthquakes. Although the p-value for the lateral spread model suggest that the existing Hazus function may be acceptable for application in Christchurch, it is proposed to update the function with the observed data anyway so that it is more specific to Christchurch. Table 4.39 summarises the new fragility function parameters obtained from this procedure and the corresponding

predictions, although since the test and training datasets are the same, this is presented only for information, not validation. It is noted that whilst the Bayesian updating procedure has significantly reduced the median value for the lateral spread function (from 1.52 m), this has not been sufficient to alter the predictions.

Table 4.39 – New parameters for liquefaction fragility functions from Bayesian updating of Hazus functions with corresponding predictions

Model	**Median PGDf**	**Dispersion**	**Predicted**		**Observed**	
			Failure	**No failure**	**Failure**	**No failure**
Vertical settlement	0.48	1.18	9	52	5	56
Lateral spread	1.03	1.54	0	5	3	2

4.6 Step 5 – Damage assignment

The first step in damage assignment is the application of fragility or repair rate functions to determine the failure probabilities of the asset with respect to each earthquake hazard (ground shaking and liquefaction). To apply the fragility function, the relevant IM and ground deformation values must be assigned to the asset. For nodes, i.e. substations, wells and pumping stations, values are assigned to an asset corresponding to the analysis grid cell (see Figure 4.16) in which it the asset located. The relevant fragility functions for each hazard are applied to estimate the failure probabilities. For linear assets, i.e. buried cables and pipes, damage assignment is less straightforward. Since linear assets can cross cells, in some cases it is possible to attribute multiple values of intensity to each individual asset. As discussed in section 4.5.2, it is preferable in these cases to discretise the asset according to hazard intensity rather than assigning a single value. Consequently cables and pipes with this situation are divided at cell borders and the corresponding hazard intensity is assigned to each section. The relevant repair rate function is applied to estimate the mean repair rate for each section, RR_{MEAN}, and this is converted to a failure probability by applying Eq. 4.69. For an asset made up of *j* sections, the failure or damage probability for the asset as a whole is given by Eq. 4.70.

$$P\left[\, fail\ or\ damage \right] = 1 - e^{-RR\left\{L_{cable};L_{pipe}\right\}} \tag{4.69}$$

$$P\left[\, fail\ or\ damage \right] = 1 - \prod_{i=1}^{j} P\left[no\ fail\ or\ no\ damage \right]_i \tag{4.70}$$

The application of a fragility function to an asset only evaluates the failure probability with respect to a single earthquake hazard. However, in events where there is more than one earthquake hazard, e.g. due to liquefaction, the overall failure probability of the asset must account for the failure probability due to all hazards. For all assets, the method for damage assignment proposed by Hazus (FEMA, 2015) is adopted. The overall failure probability for an asset is given by Eq. 4.71

$$P\left[\, fail \right] = 1 - P\left[no\ fail \right] \tag{4.71}$$

Two earthquake hazards are considered in this case study, each triggering ground shaking and liquefaction. In the Hazus method, failure due to ground shaking and failure due to liquefaction are assumed to be independent. Therefore, the probability that the asset will not fail is the product of the individual probabilities that the asset will not fail due to each hazard, as shown in Eq. 4.72.

$$P\left[\, fail \right] = 1 - P\left[no\ fail \mid PGA \right] \times P\left[no\ fail \mid PGDf \right] \tag{4.72}$$

However, there are two types of liquefaction hazard – vertical settlement and lateral spreading – and as summarised in section 4.5, the method for predicting liquefaction-induced failure probabilities varies by asset. In some cases, liquefaction-induced failure probabilities are predicted separately for vertical settlement and lateral spreading, e.g. substations, wells, pumping stations and some pipe typologies. In other cases, a single liquefaction-induced failure probability is predicted based on a composite IM such as $PGDf_{GEOM}$, e.g. buried cables and some pipe typologies. For assets that are assessed in terms of a single failure probability, the overall failure probability is given by Eq. 4.72. For assets that are assessed in terms of separate failure probabilities, Hazus proposes that the maximum of the failure probabilities is used as the

liquefaction-induced failure probability, as shown by Eq. 4.73. The rationale is that the different liquefaction hazards do not represent independent failure modes and so cannot be multiplied together, (which would overestimate the overall failure probability). Eq. 4.73 can be rewritten in terms of failure probabilities as in Eq. 4.74.

$$P\left[\, fail \,\right] = 1 - P\left[no\ fail \mid PGA \right] \times \min\left\{ P\left[no\ fail \mid PGDf_V \right]; P\left[no\ fail \mid PGDf_H \right] \right\} \tag{4.73}$$

$$P\left[\, fail \,\right] = 1 - \left(1 - P\left[\, fail \mid PGA \right]\right) \times \left(1 - \max\left\{ P\left[\, fail \mid PGDf_V \right]; P\left[\, fail \mid PGDf_H \right] \right\}\right) \tag{4.74}$$

The functionality of each asset is then assigned by randomly sampling from a uniform distribution between 0 and 1, and comparing the sampled variate to the failure probability. The asset is assumed to have failed if the variate is less than or equal to the failure probability, and is assumed to have survived otherwise. In the specific case of pipes, this method only determines whether a pipe is damaged, not whether it has failed. Therefore an additional step is needed to determine the type of damage. The assumptions of Hazus (FEMA, 2015) are adopted here. This assumes that in the presence of liquefaction, 80% of damage to pipes are breaks and 20% are leaks, with the percentages reversed in the absence of liquefaction. So for each damaged pipe, another uniform random variable is sampled between 0 and 1. If the variate is greater than 0.8, the damage is assumed to be a leak if the pipe experiences liquefaction or a break if the pipe does not experience liquefaction. If damage is classified as a break, then the pipe is assumed to have failed. If the damage is classified as a leak, then the pipe is assumed to be functional.

4.7 Step 6 – System performance

4.7.1 Electric power network

The proposed performance metric for the electric power network is the proportion of the population disconnected from the network, as discussed in section 3.2.3. Since only the sub-transmission network is analysed, the demand nodes, i.e. the downstream points at which system performance is measured, are the district/zone substations. The functionality of each

district/zone substation is determined by its physical damage state, as assigned by the method described in section 4.6, and whether there is an undamaged path between it and an operational source node, i.e. a Transpower GXP. To measure system performance it is necessary to know the population served by each district/zone substation. Orion have provided a map visualising the approximate current service area of each district/zone substation as shown in Figure 4.41.

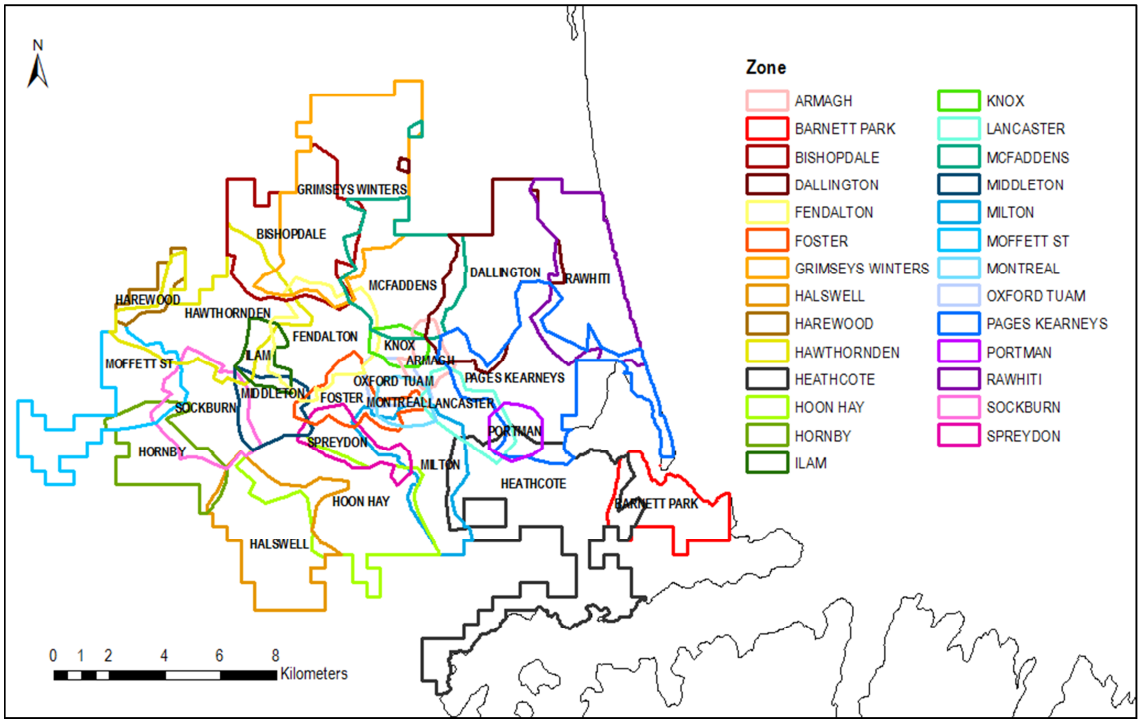

Figure 4.41 – Map of Orion district/zone substation supply areas

To use the service area map for the pre-earthquake network, some assumptions are made. The Papanui substation replaced the Harris substation after the earthquake and so it is assumed that the pre-earthquake service area for the Harris substation is the same as the current Papanui service area. The Rawhiti substation replaced the Brighton and Pages substations after the earthquake. It can be seen in Figure 4.8 that the Brighton and Pages substations are very close to each other in latitude. So for simplicity, the Rawhiti service area is divided by a horizontal line between the two substations and it is assumed that the area to the north is supplied by the Brighton substation and the area to the south is supplied by the Pages substation.

The population served within each supply area is determined using night time population data from the 2006 Census for model validation, and from the 2013 Census for future risk projection (Statistics New Zealand, 2014a). The total population in the study area is 328,814. Meshblocks are the smallest geographic unit for which statistical data are collected in New Zealand and there are 3,352 in the study area. Analysis is conducted on the grids shown in Figure 4.16. On the periphery of the grids, there some meshblocks that are only partially within the gridded area. In these cases only the part of a meshblock that is within the gridded area is analysed and so the population of the meshblock is reduced in proportion to the area of the meshblock that is within the gridded area. This affects 735 meshblocks.

The largest meshblock has a population of 963 and they have an average population of 98. It is assumed that a meshblock is supplied by a particular substation if the centroid of the meshblock falls within the supply area of that substation. In some locations, especially around the city centre, the supply areas overlap, meaning that some meshblocks can be supplied by more than one substation. After an earthquake, it is assumed that the entire population of a meshblock is connected to the power network if at least one of its supply substations remains operational. The overall system performance, SP_{EPN}, is then determined by Eq. 4.75, where Pop_i is the population of meshblock i and $w_{(i,EPN)}$ is a weight assigned to each meshblock, which equals 0 if the meshblock is connected to an operational substation and 1 otherwise.

$$SP_{EPN} = \frac{\sum_{i=1}^{3352} Pop_i \times w_{(i,EPN)}}{Total\ pop} \tag{4.75}$$

4.7.2 Water supply system

The proposed performance metric for the water supply system is the proportion of the population disconnected from the network, as discussed in section 3.2.3. The system performance SP'_{WSS} is given by Eq. 4.76, where Pop_i is the population of meshblock i and $w'_{(i,WSS)}$ is a weight assigned to each meshblock, which equals 0 if the meshblock is connected to

the water supply system and 1 otherwise. A prime symbol is included in the notation for this metric and the weights. This is to signify that this is an initial system performance value calculated before interdependencies have been accounted for.

$$SP'_{WSS} = \frac{\sum_{i=1}^{3352} Pop_i \times w'_{(i,WSS)}}{Total\ pop} \tag{4.76}$$

The method for determining whether a meshblock is connected to the water supply system is different to the case of electric power network. The use of a simplified network means that only major pipes and associated nodes are included in the analysis. Therefore, the only obvious demand nodes in this simplified network described in section 4.2.2 are the 41 booster pumping station sites. There are no defined service areas for these assets so they must be determined manually. The extent of the supply area of each booster pumping station is determined visually by overlaying the pipes shapefile onto the map of meshblocks and satellite imagery for the elevated parts of the city. Two methods are then used to assign each meshblock to the supply area of a pumping station. If a meshblock contains a pipe that is downstream of a pumping station, then it is assumed to be part of the supply area of that pump. If a meshblock does not contain a pipe that is downstream of a booster pumping station, but is in a part of the city that is expected to receive water by pumping (e.g. the south and southeast due to the Port Hills), then based on the assumption that pipes are generally constructed along roads, the satellite imagery is examined to determine which booster pumping station is the most likely source for that community.

However, most of Christchurch is not served by booster pumps and so to represent these areas, demand nodes must be manually created. Creating a demand node at the end of every pipe would lead to a very large number of demand nodes, including some that are very close together. Not only does this increase computation time, it also leads to some meshblocks containing more than one demand node and so assumptions would have to be made regarding how populations are assigned to each one. It is proposed to create demand nodes at less regular

intervals along the pipe network. Initially demand nodes are created at all simplified network endpoints and also at all junctions where paths from two different wells meet. Then, to increase the resolution of the analysis, additional demand nodes are located subjectively at intermediate pipe junctions based on factors including population density and meshblock size. This is done to ensure that there are more demand nodes in densely developed areas and that no meshblock contains more than one demand node. A total of 299 manually assigned demand nodes, referred to as sinks, are created in addition to the 41 booster pumping station sites, as shown in Figure 4.42. The average separation distance of the manual demand nodes is 571m.

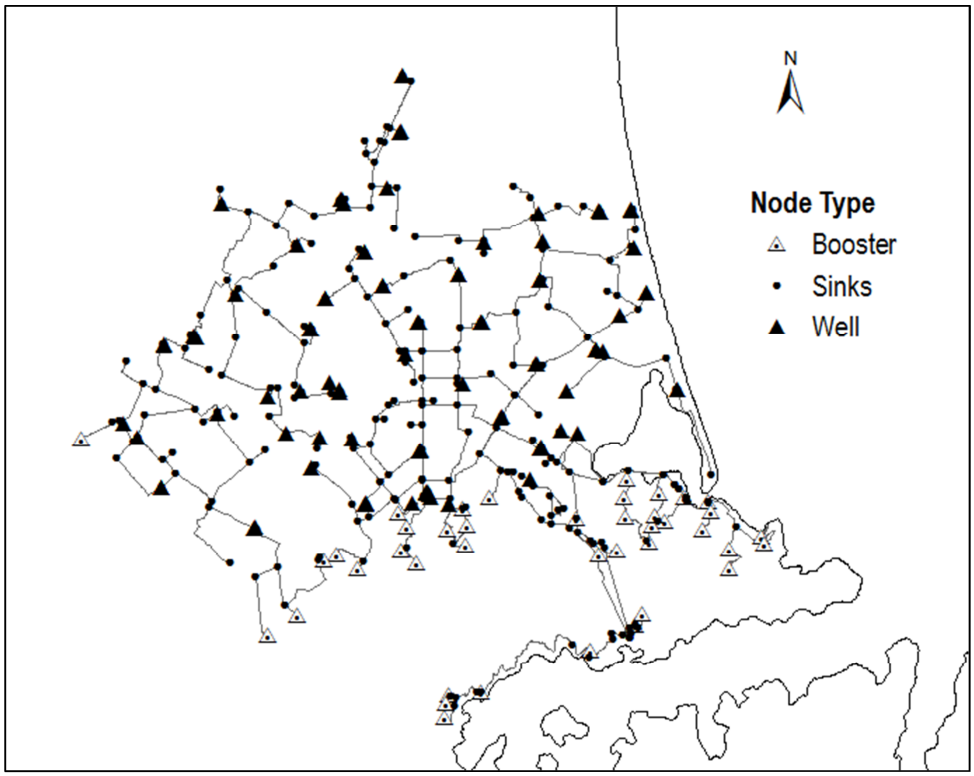

Figure 4.42 – Map showing location of nodes in water system performance analysis, including wells (primary pumping stations), booster pumping station sites and demand sinks

Each meshblock is assigned to either a booster pumping station, if the meshblock falls within its supply area, or otherwise it is assigned to the closest sink. Each meshblock is only assigned to a single sink. The entire population of a meshblock is deemed to be connected to the water network if there is an undamaged path to its assigned sink node from any operational well

within its pressure zone. Where a meshblock is assigned to a booster pumping station sink node, the pump must also itself be operational.

4.7.3 Interdependency simulation

In addition to calculating the performance of each system independently, the interdependent performance of the water supply system must be calculated as this represents its true performance. The choice of interdependency model is dependent on the type of data available. For this case study there is sufficient information to link assets across systems, and so an integrated model is proposed in the form of a nodal analysis, as described in section 2.6.2. The interdependency between the two systems exists because pumping stations rely on electric power. This includes booster pumping stations and wells, since they are attached to primary pumping stations. Two methods are proposed for linking the two assets: supply relationship and geographical proximity.

Each pumping station is located in a specific meshblock, and each meshblock is assigned to one or more district/zone substations. In this way it is possible to determine a supply relationship between each pumping station and the district/zone substations. This method allows for pumping stations to be connected to more than one substation and so accounts for redundancy in power supply. An alternative method, adopted in some other studies (e.g. Pitilakis et al., 2014a), is to assign each dependent asset (i.e. pumping station) to its closest supporting asset (i.e. substation). This is a simplification that does not account for redundancy and so may potentially be conservative. Both methods are tested in this study to assess the validity of the geographical proximity assumption.

For both methods, if a district/zone substation is designated as failed, then this failure is propagated across the identified interdependency links to downstream dependent assets. A specific pumping station (and corresponding well if applicable) is designated as failed if all of its supply substations have failed and it does not have a back-up power supply. The final interdependent performance of the water supply system, SP_{WSS}, is then calculated for each

linkage method, by re-applying Eq. 4.76 with a new set of weights, $w_{(i,WSS)}$, that account for the new operational states.

4.8 Model validation

The Christchurch seismic risk assessment model is initially tested against observations from the Christchurch earthquake. As discussed in Chapter 1, this serves both as an example of model application for scenario events (for either rapid assessment or regulatory purposes) and in a limited capacity as a model validation exercise. It is also proposed to use the model validation to test three aspects of the model methodology: the analysis grid resolution, which is either 1 km or 500 m as in Figure 4.16; the choice of liquefaction triggering model, which is either model LPI1 (LPI based on shear-wave velocity profile) or LPI3b (LPI based on simulated shear-wave velocity profile) as discussed in section 4.4.6; and the application of interdependency to the water supply system (either by substation zone or by substation proximity) as discussed in section 2.6.2. This means that eight different versions of the Christchurch model are tested in the validation exercise, as shown in Table 4.40. Since the method chosen for applying interdependency only affects water supply system performance, the 'a' and 'b' versions of each method produce identical results for the electric power network.

Table 4.40 – Versions of Christchurch model to be validated against observations from Christchurch earthquake

Model	Grid resolution	Liquefaction triggering model	Interdependency
Version 1a	1km	LPI1	Substation zone
Version 1b	1km	LPI1	Substation proximity
Version 2a	500m	LPI1	Substation zone
Version 2b	500m	LPI1	Substation proximity
Version 3a	1km	LPI3b	Substation zone
Version 3b	1km	LPI3b	Substation proximity
Version 4a	500m	LPI3b	Substation zone
Version 4b	500m	LPI3b	Substation proximity

The GMPE adopted in the model has already been validated against the Christchurch earthquake (Bradley, 2013). Therefore the observed PGA map (Figure 4.1), produced from that GMPE (Bradley, 2013) is used as input for the validation, along with the corresponding PGV ShakeMap (Figure 4.2). In the case of PGA, the map produced for the New Zealand Geotechnical Database by Bradley and Hughes (2012) does not quite fill the extent of the proposed analysis grids. In the 1 km grid, 34 of the 294 grid locations do not have a PGA value and in the 500 m grid, 77 of the 1,007 grid locations do not have a PGA value. The majority of the affected grid locations are in the Port Hills area to the south of the city. USGS ShakeMap PGA observations are available across the entire grid area (USGS, 2015b) and could be used as an alternative. However, there are significant discrepancies between the NZGD and USGS datasets. When comparing PGA observations at grid locations for which both data sources are available, USGS observations are on average over 30% larger and so the two cannot simply be used interchangeably. It is assumed here that the NZGD dataset is more reliable since it has been produced locally and the specific methodology used to develop the maps, including local strong motion station records and locally specific ground motion prediction equation and spatial correlation model for interpolation, has been described in detail by Bradley and Hughes (2012). Conversely, there is more limited information available on the production of the ShakeMap for the Christchurch earthquake (USGS, 2015b). It is proposed therefore to use the NZGD PGA maps but to supplement this with USGS data at locations outside the NZGD map extents. Since the USGS data appear to overestimate PGA, it is proposed to use the data at locations where both NZGD and USGS data are available to develop predictive relationships between the two datasets so that a 'NZGD equivalent' value of PGA can be estimated from the USGS value.

Figure 4.43 shows plots of the corresponding PGA values from each dataset, separately for the 1 km grid and the 500 m grid. Both cases show a general linear trend and ordinary least squares regression is used to derive a linear predictive function. Since both cases yield high coefficients of determination (R^2) and p-values confirm that the functions are statistically significant at the

5% level, it is reasonable to use these functions to simulate 'NZGD equivalent' PGAs at the non-mapped grid locations.

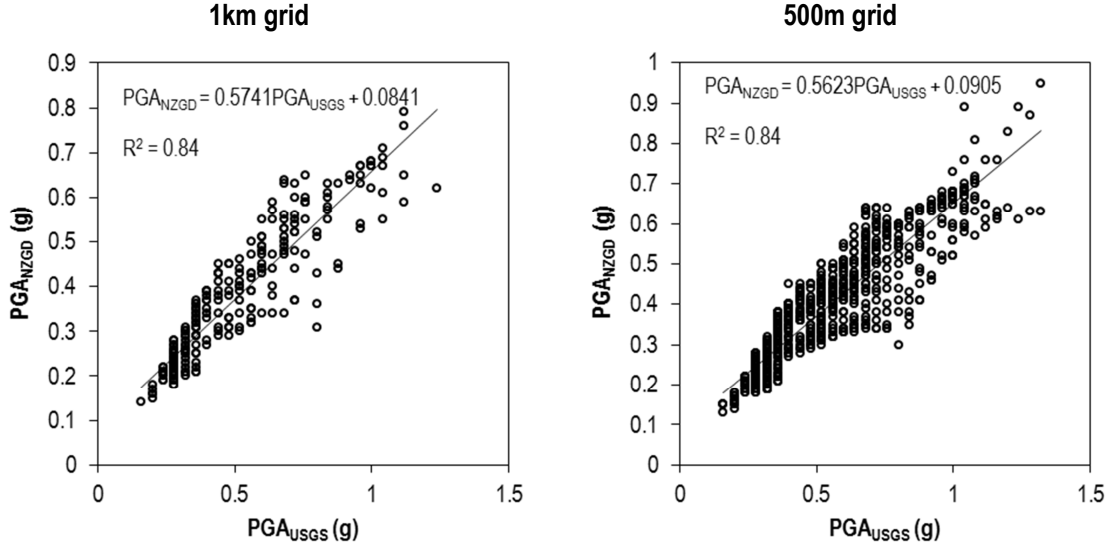

Figure 4.43 – Linear relationships linking PGA observations from NZGD dataset to PGA observations from USGS dataset

In order to account for uncertainties in individual model components, the model is run 1,000 times in a Monte Carlo simulation with re-sampled random variables. Sources of uncertainty include, ground motion prediction, prediction of liquefaction-induced permanent ground deformations and the application of fragility and repair rate functions to assets, all of which require an element of random sampling. With respect to PGA, the uncertainty is small and has been quantified by Bradley (2013) as part of the GMPE validation. The uncertainty is assumed to be the same across the affected area and since the same PGA and PGV inputs are used for all simulations, the effects of uncertainty from other model components can be more closely investigated without being masked by the influence of uncertainty from ground motion. The following sections present the results of the Monte Carlo simulation and a geospatial validation of the results.

4.8.1 Simulation results

Figure 4.44 to Figure 4.46 show histograms of the system performance results obtained when each version of the model listed in Table 4.40 is run. The performance of the system is

measured as the percentage of the study area population disconnected from the system, (rounded to the nearest 1%). The histogram displays the distribution of results across the 1,000 simulations for each version of the model. In addition, the observed performances are displayed on each plot, (71% for the electric power network and 80% for the water supply network as summarised in section 3.2), as well as the mean value from all simulations.

Figure 4.44 shows the results for the electric power network. It can be seen that all versions of the model generate performance distributions that are predominantly better than observed (i.e. fewer customers disconnected) but versions 3 and 4 produce results closer to those observed than versions 1 and 2. Model version 1 only produces nine simulations out of 1,000 with worse performance than observed, i.e. if version 1 is assumed to be applicable, the probability of the system performing at least as badly as was observed is just 0.013, indicating that this model version has very low applicability to Christchurch. Version 2 can be considered not to be applicable at all since no simulations produced a performance as extreme as the observed value. For version 3, the probability of a performance equalling or exceeding the observed value or worse is 0.223 and for version 4 this figure is 0.180. Although these probabilities are relatively low, they show that the model could plausibly simulate scenarios similar to the Christchurch earthquake. In both cases, the observed performance is within one standard deviation of the mean model prediction. For version 3, the absolute error is 5.3%, which is equivalent to a relative error of 7.5% (relative to the observed performance). This compares favourably to previous studies that have attempted to validate infrastructure system performance model with real natural hazard events (Wang and O'Rourke, 2006; Javanbarg and Takada, 2009; Winkler et al., 2010; Nateghi et al., 2011). Thus versions 3 and 4 are deemed applicable for the case of the Christchurch electric power network. Versions 1 and 2 use model LPI1 for liquefaction triggering while versions 3 and 4 use model LPI3b. Although the analysis in section 4.4.6 shows that both of these models have a similar predictive capability, it is possible that there is a geographic bias in each model such that LPI1 is poorer at predicting liquefaction occurrence in

areas that are more densely populated and contain a higher concentration of network assets, and hence under-predicts the damage to, and performance of, the system.

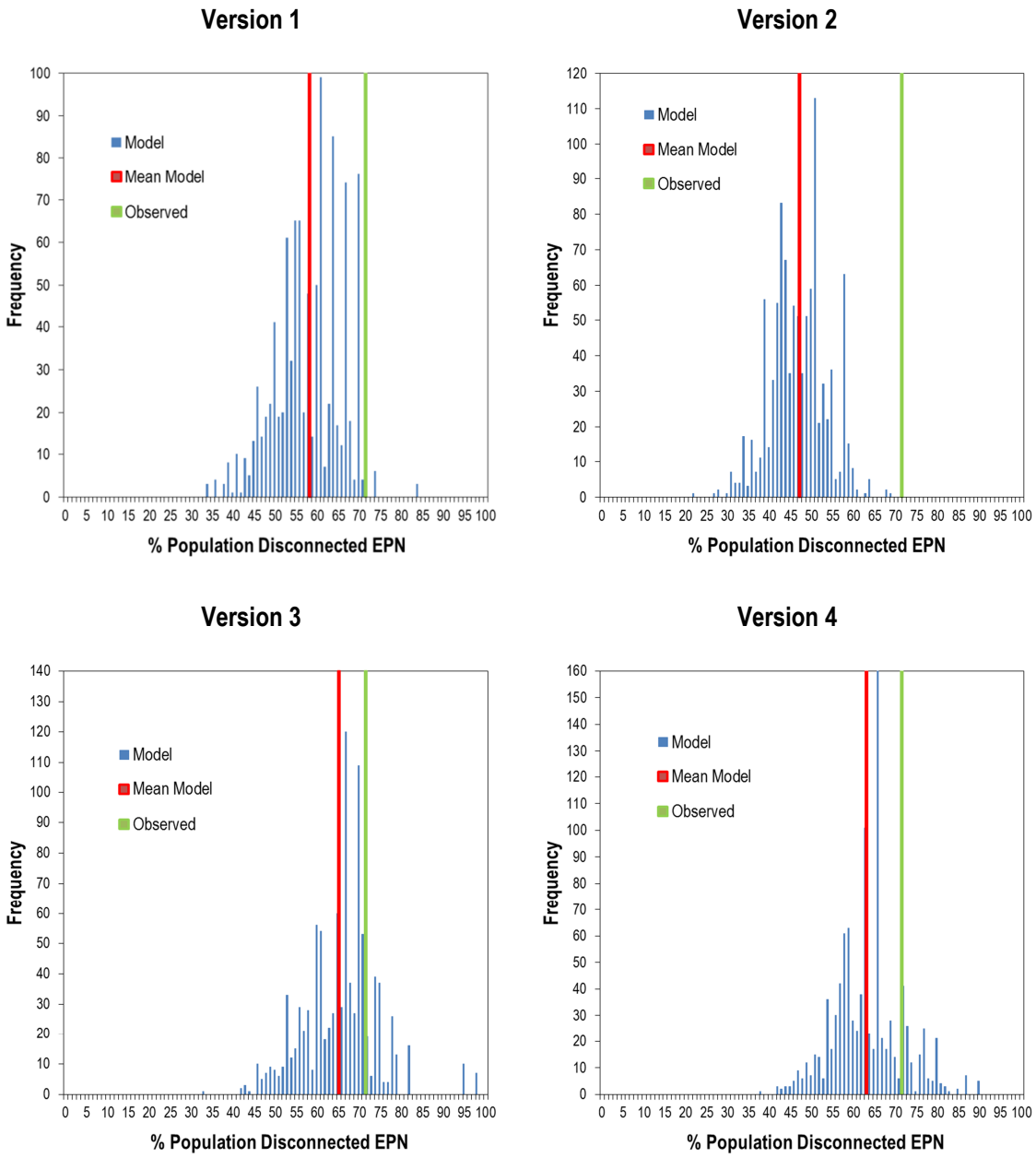

Figure 4.44 – Results of electric power network performance from model runs with observed Christchurch earthquake ground motion

Figure 4.45 shows the results for the water supply system with interdependency applied by substation zone analysis. A notable observation is that the range of the results is much narrower and more symmetrical than for the electric power network. This is because there are significantly more demand nodes in the water supply system than in the electric power network

and so the analysis is at a higher resolution in population terms. In the electric power network there are 30 demand nodes and so on average each demand node serves 3.33% of the population. Therefore for two simulations, in which the only difference is the operational state of one demand node, the average change in performance is 3.33%. In the water supply system there are 340 demand nodes and so the corresponding average change in performance would be 0.29%.

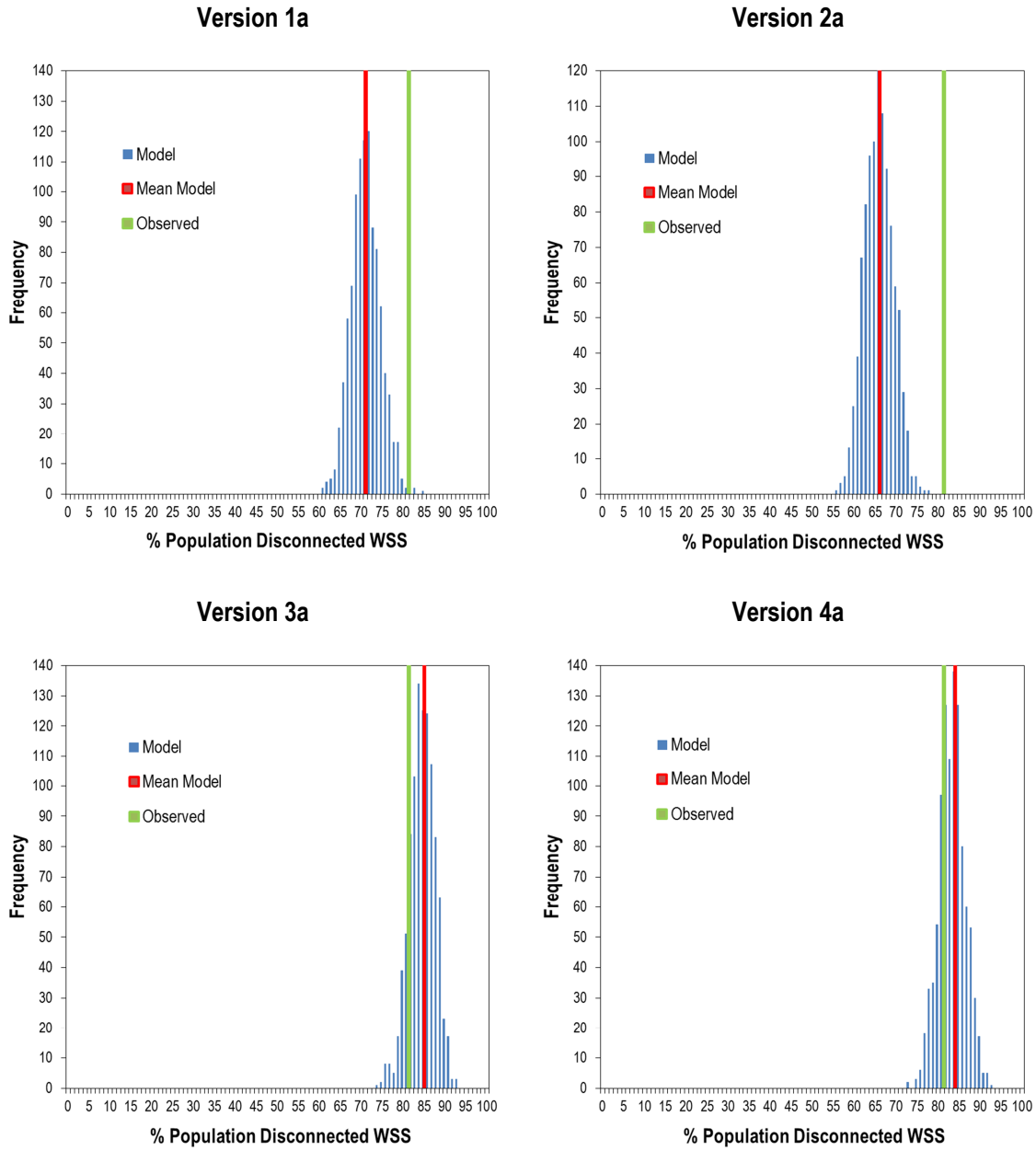

Figure 4.45 – Results of water supply system performance (with interdependency by zone) from model runs with observed Christchurch earthquake ground motion

It can be seen that, as was the case with the electric power network, versions 3a and 4a produce results closer to those observed than versions 1a and 2a, although in this case the results for versions 3a and 4a are mainly worse than observed. For model version 1a, the histogram shows that the probability of a performance at least as poor as observed is just 0.005. For version 2a there are no simulations that are as extreme as the observed value. For version 3a, the probability of a performance better than or equal to observed value is 0.132 and for version 4 this figure is 0.248. As with the electric power network, these versions of the model imply that the Christchurch earthquake and its consequences on the water system are rare events but ones that can be plausibly simulated. The observation is within one standard deviation of the mean prediction from version 4a, but is 1.26 times the standard deviation away from the mean prediction from version 3a. However, the relative error of the mean predictions are 4.8% for version 3a and 3.2% for version 4a, both of which compare favourably to previous studies (Wang and O'Rourke, 2006; Javanbarg and Takada, 2009; Winkler et al., 2010; Nateghi et al., 2011). This validates versions 3a and 4a for application to the Christchurch water supply system, whilst versions 1a and 2a can be discarded.

Figure 4.46 shows the results for the water supply system with interdependency applied by substation proximity. It can be seen that in this case, versions 1b and 2b produce results closer to those observed than versions 3b and 4b. The results for versions 1b and 2b show that they mainly predict performances better than observed. Instead, versions 3b and 4b result in predictions that are mainly worse than observed. For model version 1b, the histogram shows that the probability of a performance at least as poor as observed is 0.121 and for version 2b this figure is 0.21. For version 3b, the probability of a performance better than or equal to the observed value is 0.007 and for version 4b this figure is 0.008. The observation is within one standard deviation of the mean prediction from version 2b, but is 1.29 times the standard deviation away from the mean prediction from version 1b. However, the relative error of the mean predictions are 5.1% for version 1b and 3.9% for version 4b, both of which compare favourably to previous studies (Wang and O'Rourke, 2006; Javanbarg and Takada, 2009;

Winkler et al., 2010; Nateghi et al., 2011). This validates versions 1b and 2b for application in Christchurch for the water supply system, whilst versions 3b and 4b can be discarded.

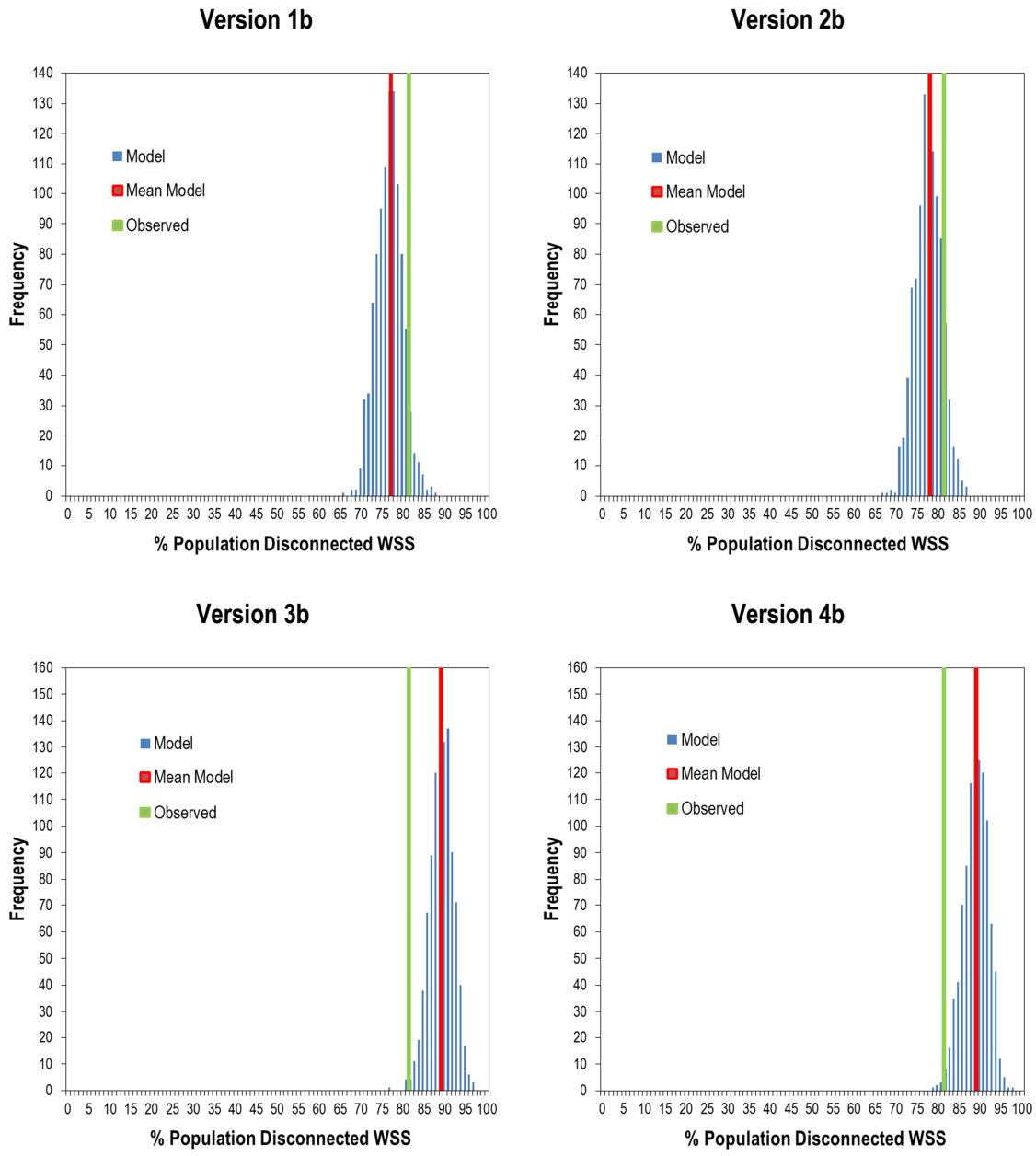

Figure 4.46 – Results of water supply system performance (with interdependency by proximity) from model runs with observed Christchurch earthquake ground motion

Three model parameters have been tested in the validation exercise: interdependency methodology, liquefaction triggering model, and analysis grid resolution. The relative performance of the two interdependency methodologies is considered by comparing the 'a' versions of the model (interdependency by substation zone analysis) to the corresponding 'b'

version (interdependency by proximity). It is noted that in each case, performances are on average worse when interdependency is analysed by proximity. This result is anticipated since when interdependency is analysed by proximity, redundancy is not taken into account. When interdependency is analysed by substation zones, each pumping station (both primary and booster) can receive power from one or more substations as per the network topology. The pumping stations are therefore more resilient since even if one substation fails, the pumping station may still be able to function, thus improving the overall system performance. Conversely when interdependency is analysed by proximity, each pumping station is assigned to a single substation. If that substation fails, the pumping station also fails, and so overall system performances are worse. In general one might expect that because it does not account for the real-world redundancy, the proximity method would always over-predict damage but it is noted that in versions 1b and 2b, the simulation predictions are generally better than observed. Both of these versions use liquefaction triggering model LPI1 and so there is an implication that LPI1 may under-predict damage. At this point the choice between the two methods is inconclusive since there are validated model versions using both the substation zone method (3a and 4a) and the proximity method (1b and 2b).

To consider the relative impacts of the alternative liquefaction triggering models, the remaining parameters must be kept constant so version 1 is compared to version 3 and version 2 is compared to version 4. In all cases the model versions with liquefaction triggering model LPI3b (i.e. model versions 3 and 4) on average simulate worse performances than the corresponding model versions with liquefaction triggering model LPI1 (i.e. model versions 1 and 2). This supports the notion that model LPI1 under-predicts damage. Although the analysis in section 4.2 shows that overall both models have a similar predictive capability, it is possible that there is a geographic bias in each model such that LPI1 is poorer at predicting liquefaction occurrence in areas that are more densely populated and contain a higher concentration of network assets, and hence predict lower levels of damage and better performance of the systems. This is discussed further in section 4.8.2. For the electric power network it is evident that model versions that use

LPI3b produce results closer to the observed performance. It is essential that the final model is valid for both the electric power network and water supply system and therefore the final model must be either version 3 or version 4. Since versions 3a and 4a have been shown to be valid, but not versions 3b and 4b, then the final model must be an ‘a’ model, i.e. it must use the substation zone method for interdependency.

Grid resolution is considered by comparing model version 3a (1 km) to version 4a (500 m). For both the electric power network and water supply system, predicted performances are more conservative (i.e. higher proportion of disconnections) when the 1 km grid model is used. In the case of the electric power network, the consequence is that the mean prediction is closer to the observed performance when the 1km grid is used. Instead, for the water supply system, the mean prediction is closer to the observed performance when the 500 m grid is used. Intuitively one would expect a higher grid resolution to enhance the predictive power for a model. The fact that this is not the case for the electric power network, coupled with the fact that the 1 km grid model still provides reasonable, albeit slightly more conservative, predictions for the water supply system, brings into question whether the additional work and time associated with the 500 m grid is worthwhile. For the 1,000 simulations in this validation exercise, the 1 km grid model had an average run time of 15.3 seconds per event and the 500 m grid had an average run time of 23.5 seconds per event, more than 50% longer, (on a Windows PC with 2 GB of RAM and a 2.5 GHz processor). If this is extrapolated to a full 10,000 year seismic risk assessment with a rate of 1.45 events per year (see section 4.3.2), then this corresponds to a run time of around 62 hours for the 1 km grid and around 94.5 hours for the 500 m grid. On balance, the benefits of using the higher resolution grid are inconclusive and when set against the additional time costs incurred, there is little evidence to support the adoption of the 500 m grid. Therefore, the 1 km grid, i.e. version 3a, is adopted for the final model. This is in line with the analysis resolution adopted by Hazus (Neighbors et al., 2013), SYNER-G (Pitilakis et al., 2014a) and RiskScape (NIWA, 2015).

4.8.2 Geospatial validation

In addition to checking how well the proposed final model (version 3a) predicts overall performance, it is also instructive to compare how well the model predicts the spatial distribution of performance across the city. For the electric power network this is relatively straightforward since the observed functionality (i.e. operational or failed) for each meshblock can be deduced by overlaying the meshblocks onto the observed outage map in Figure 4.9. A meshblock is then assumed to have failed if its centroid is within the mapped outage area. Cross-referencing with the outage map identifies that 2,402 meshblocks lost electric power supply and 950 remained connected following the Christchurch earthquake. To compare this spatially with the model predictions, contingency table analysis is applied in the same manner as for the validation of liquefaction triggering models in section 4.4.6. For each meshblock there are 1,000 models simulations of the electric power network functionality. The average model prediction for that meshblock is simply the binary state (operational or failed) that is predicted most frequently in the simulations. For the purposes of this validation, a positive prediction is defined as the average model prediction of the meshblock being operational and therefore, in terms of a contingency table analysis, this is equivalent to a predictive model for the operability of a meshblock with a threshold of 500. Table 4.41 summarises the contingency table data comparing the average model predictions for each meshblock to the observed state and the corresponding diagnostic scores: true positives (TP), true negatives (TN), false positives (FP), false negatives (FN), true positive rate (TPR, the ratio of true positive predictions to observed positives), true negative rate (TNR, the ratio of true negative predictions to observed negatives), false positive rate (FPR, the ratio of false positive predictions to true negatives) and false negative rate (FNR, ratio of false negatives to true positives). In total the model correctly predicts the functionality at 2,393 meshblocks (71%). A good model is represented by high scores (>0.5) for TPR and TNR and low scores (<0.5) for FPR and FNR. Table 4.41 shows that these conditions are met by the model and therefore it can be concluded that model satisfactorily replicates the geospatial distribution of electric power network functionality across meshblocks.

Table 4.41 – Contingency table results and diagnostic score for average model prediction for electric power network functionality

TP	TN	FP	FN	TPR	TNR	FPR	FNR
503	1890	512	447	0.529	0.787	0.213	0.471

For the water supply system a geospatial validation of functionality is not possible since there is no information on how observed system outages were distributed across the city. As an alternative it is proposed to assess the geospatial validity of the model by comparing the predicted and observed proportions of pipe repairs in each of the pressure zones (Figure 4.13) in the city in a chi-square goodness of fit test (Hines et al., 2003). The input data for the chi-square test are shown in Table 4.42.

Table 4.42 – Observed and predicted distribution of pipe repairs across pressure zones

Zone	Expected proportion	Expected frequency	Observed frequency
Central	76%	149	152
North West	7%	13	10
West	3%	7	8
Riccarton	1%	2	5
Rocky Point	9%	18	15
Parklands	4%	7	6
Total		196	196

The expected proportions of pipe repairs in each zone are taken as the mean value from the 1,000 model simulations. To enable direct comparison between model predictions and observations, only the simplified pipe network, identified in section 4.2.2, is used in the comparison. A total of 196 repairs were observed in this part of the network after the Christchurch earthquake. The null hypothesis of the chi-square test is that the proportions of repairs in each zone are as expected from the model simulations and the alternative hypothesis is that at least one of the proportions is different. The dataset has five degrees of freedom and a

significance level of 0.05 is used for the analysis. Based on the observed and expected frequencies, the chi-square test statistic is calculated to be 6.04. The p-value is the probability of the chi-square test statistic having five degrees of freedom being more extreme than 6.04 and is calculated to be 0.30. Therefore at the 0.05 significance level, the null hypothesis cannot be rejected and one can conclude that the observed proportions are in line with the model predictions. The model is therefore geospatially validated for damage to the water supply system.

5. Results

5.1 Seismic risk assessment

The final validated seismic risk assessment model, (i.e. version 3a with 1km grid, liquefaction triggering model LPI3b and zonal interdependency analysis) is applied to the current (2014) electric power and water supply systems in Christchurch to assess the future risk to the city.

5.1.1 Hazard

Application of the seismic hazard model, described in section 4.3, generates 14,428 events of $M_W \geq 5$ in a 10,000 year catalogue. The frequency distribution of magnitudes in the catalogue is shown in Figure 5.1 and follows an expected pattern of lower magnitude events occurring with higher frequency. In total there are 1,801 event with $M_W \geq 6$, which is just over 1 every 5.5 years and compares to just under 1 every 5 years in the historic catalogue. There are 520 events with $M_W \geq 7$, which is approximately 1 every 19 years and compares to 1 every 25 years in the historic catalogue. There are 29 events with $M_W \geq 8$, which is approximately 1 every 345 years.

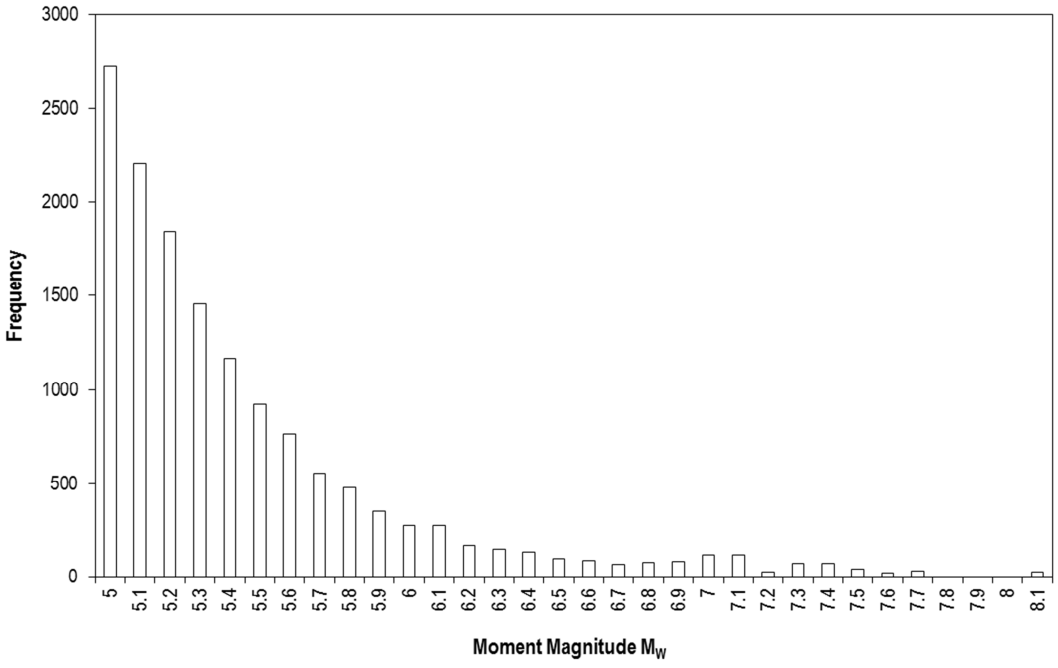

Figure 5.1 – Frequency distribution of earthquake magnitudes in simulated 10,000 year catalogue with a total of 14,428 events

It is common in loss estimation to report metrics on an annualised basis and Figure 5.2 shows the frequency distribution of the maximum annual magnitudes, i.e. the highest magnitude event observed to occur in each of the 10,000 years of the catalogue. A notable observation is that in almost a quarter of the catalogue (2,278 years), no earthquakes of $M_W \geq 5$ are observed to occur. Since many of the low magnitude events will occur in years when other larger events also occur, the shape of the graph in Figure 5.2 differs from Figure 5.1, with a significant reduction in the absolute frequency of low magnitude events, although in relative terms they still occur more frequently than high magnitude events.

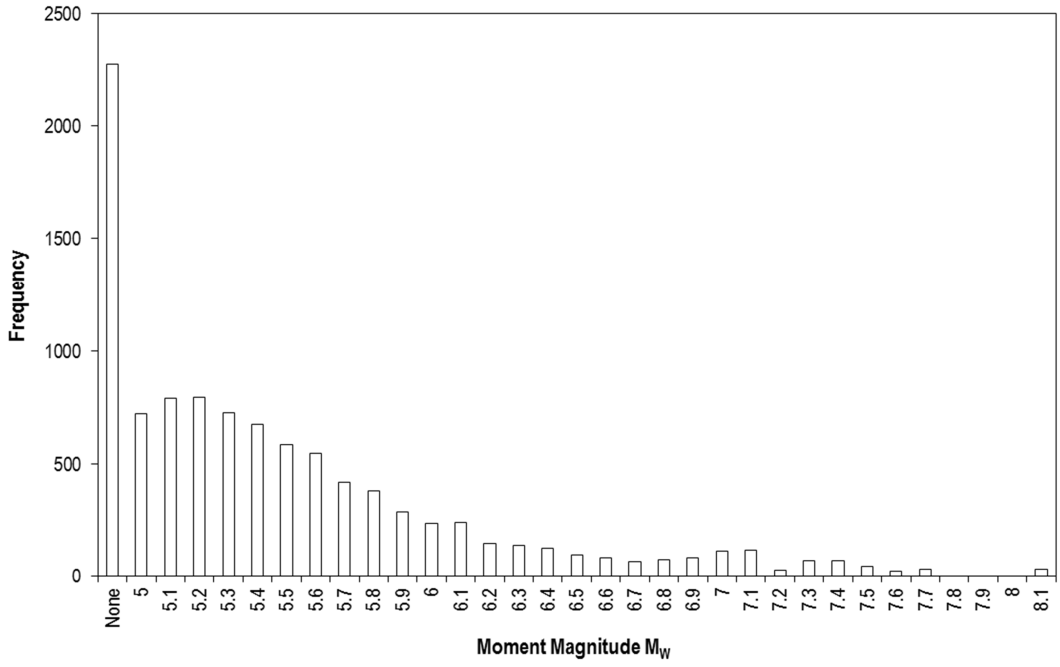

Figure 5.2 – Frequency distribution of maximum annual earthquake magnitudes in simulated 10,000 year catalogue

Figure 5.3 shows the geospatial distribution of events in the simulated catalogue with $M_W \geq 6$ (for clarity). They are generally evenly spread across the 200km buffer zone although with a denser cluster in the quadrant to the north west of Christchurch around the Alpine Fault. It should be noted that the points plotted on Figure 5.3 represent the centroid of the earthquake source. In the case of active faults, this means that multiple events generated by the same source are plotted in the same location. Hence, even though there are 29 $M_W \geq 8$ events in the

catalogue, there is only one location visible in Figure 5.3, since they are all generated by the same fault (the Fjord-Kelly section of the Alpine Fault).

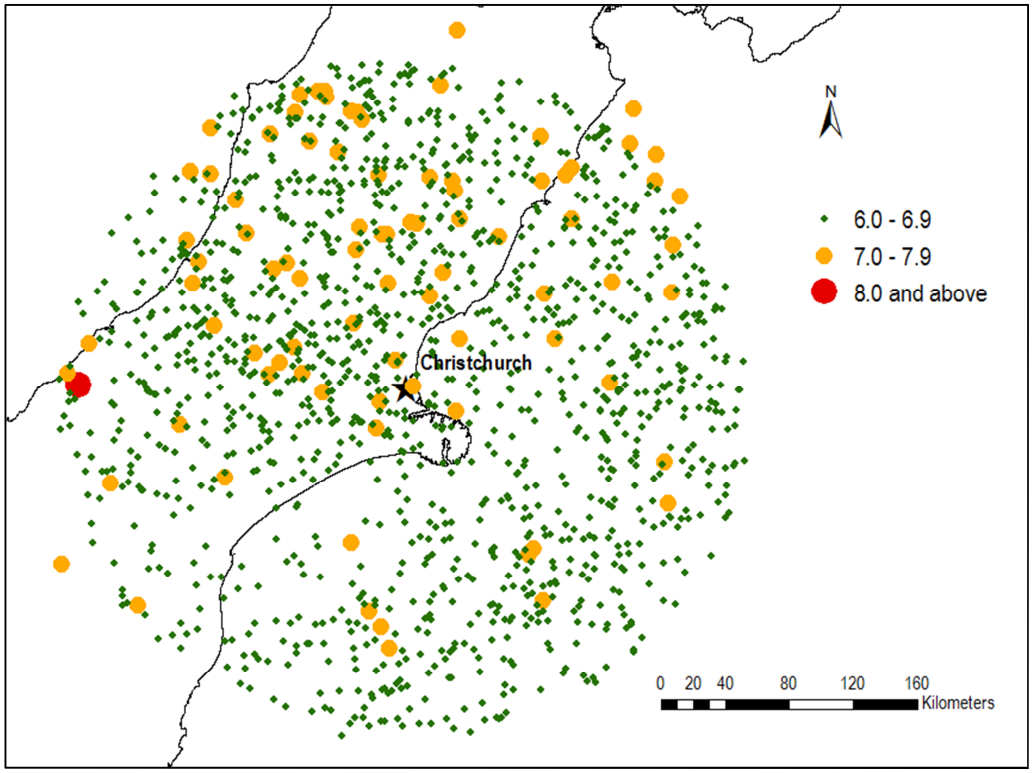

Figure 5.3 – Geospatial distribution of events in simulated 10,000 year catalogue with $M_W \geq 6$

Table 5.1 summarises the frequency distribution of the maximum ground motions observed in the study area for each event. For comparison, in the Darfield and Christchurch earthquakes, the maximum PGAs observed are greater than 0.4 g and 0.9 g, respectively. These values are exceeded 244 (once every 41 years) and 42 times (once every 238 years) respectively in the simulated catalogue. In terms of PGV, the maxima observed in the Darfield and Christchurch earthquakes are greater than 31 cm/s and 80 cm/s, respectively. In the simulated catalogue these values are exceeded 585 times (once every 17 years) and 75 times (once every 133 years), respectively.

Table 5.1 – Frequency distribution of maximum ground motions per event as observed in study area

Max PGA (g)		Frequency	Max PGV (cm/s)		Frequency
< PGA	PGA ≤		< PGV	PGV ≤	
	0.1	12702		10	11894
0.1	0.2	1006	10	20	1399
0.2	0.3	318	20	30	560
0.3	0.4	158	30	40	263
0.4	0.5	84	40	50	114
0.5	0.6	49	50	60	59
0.6	0.7	33	60	70	38
0.7	0.8	22	70	80	26
0.8	0.9	14	80	90	13
0.9	1.0	11	90	100	11
1.0	1.5	26	100	150	32
1.5		5	150		19

Table 5.2 summarises the extent of liquefaction in each simulated event in the catalogue in terms of the number of grid squares affected (out of a total of 294), and shows that liquefaction is only predicted to occur in 882 events (approximately 6%). Of these, 68 events are seen to affect over 150 sites in the study area (approximately half the city), with 29 events observed to affect over 200 sites (approximately two thirds of the city). In the case of two events, over 250 sites in the city are affected. The latter events are a M_W 7 and a M_W 6.7 event, both of which have hypocentres located directly under the city and result in 254 and 252 sites being affected by liquefaction, respectively.

Table 5.2 – Summary of number of analysis grid sites affected by liquefaction in each event

No. of sites affected	Frequency
0	13546
1 – 10	395
11 – 20	105
21 – 30	69
31 – 40	44
41 – 50	38
51 - 60	26
61 – 70	24
71 – 80	15
81 – 90	15
91 - 100	16
101 - 150	69
151 – 200	37
201 – 250	27
> 250	2

Figure 5.4 presents the liquefaction occurrence data by site, with the colours representing the number of times each grid square experiences liquefaction within the simulated catalogue. The maximum liquefaction occurrence frequency is 290 events and as expected, the most frequently affected sites (dark orange and red) are in the central and north eastern parts of the city, following the path of the River Avon where more liquefiable soil types are found.. Similarly, the least frequently affected areas (dark green) are to the south of the city in the Port Hills area where the ground is more rocky. Here there are two sites that are completely unaffected by liquefaction and 24 sites that are only affected by one event.

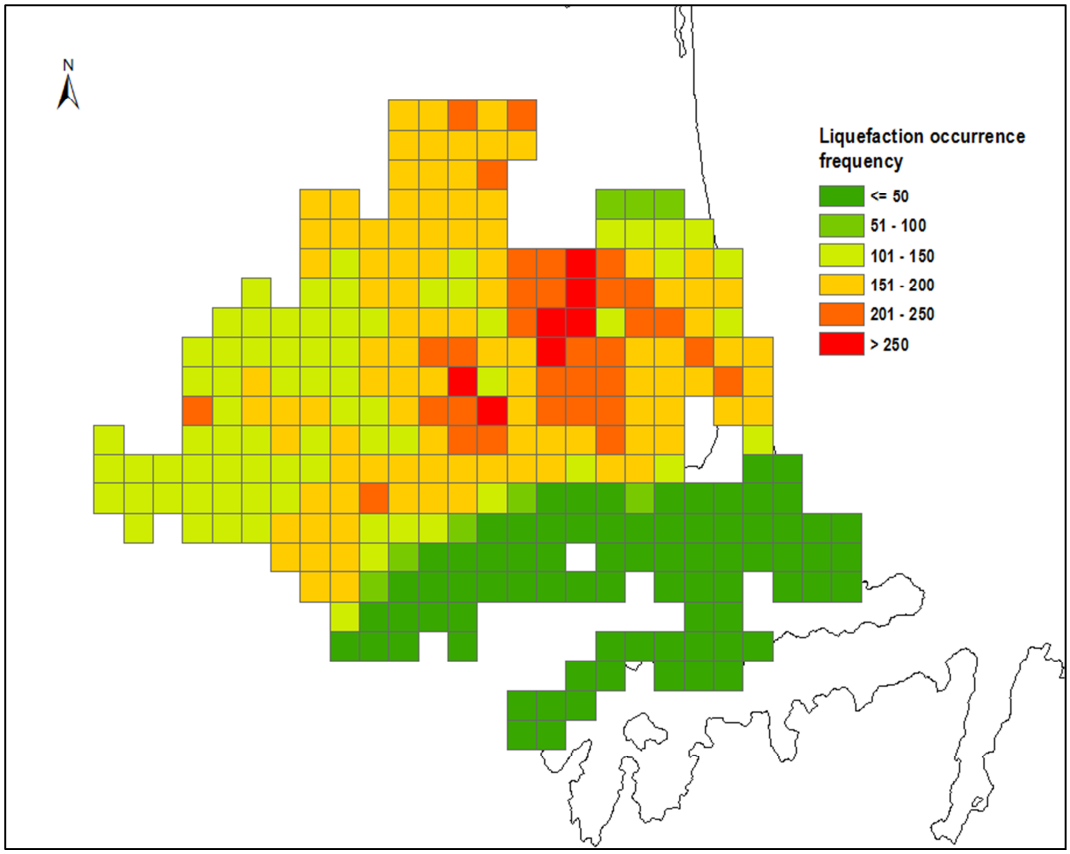

Figure 5.4 – Frequency of liquefaction occurrence within simulated catalogue at each grid square in study area

5.1.2 Electric power network

Table 5.3 and Table 5.4 summarise the effect of each event in the simulated catalogue on electric power network assets in terms of the number of failed cables (from a total of 244) and substations (from a total of 32). This shows that the vast of the majority of events do not affect the functionality of any assets, with cables failures only observed in 200 events (1.4% of all events) and substation failures only observed in 224 events (1.6%). The maximum number of cables affected in a single event is 72 (30%), caused by the same M_W 6.7 event that is seen to lead to 252 grid squares experiencing liquefaction. The same event is also associated with the occurrence of the maximum number of substation failures – 24 (75%). These low failure frequencies are in agreement with the findings of the fragility analysis conducted in Chapter 4, where the repair rate function analysis (section 4.5.2) shows that buried cables are primarily affected by liquefaction-induced permanent ground deformations. The results in Table 5.2 show

that this occurs only in a small number of events and in many of these, liquefaction may not occur directly under a cable or may not induce a sufficiently large permanent ground deformation to cause failure. The fragility function analysis in section 4.5.1 shows that the median PGA for causing substation failure is very high (1.13 g), reflecting the seismic strengthening program implemented by Orion in the 1990s (Giovinazzi et al., 2011).

Table 5.3 – Frequency distribution of number cable failures per event in simulated catalogue

No. of cables failed	Frequency
0	14228
1 – 10	130
11 – 20	33
21 – 30	14
31 – 40	13
41 – 50	3
51 - 60	2
61 – 70	3
> 70	2

Table 5.4 – Frequency distribution of number of substation failures per event in simulated catalogue

No. of substations failed	Frequency
0	14204
1 – 5	175
6 – 10	32
11 – 15	12
16 – 20	3
> 20	2

Figure 5.5 presents the failure occurrence frequency data by asset, with the colours representing the number of times each asset fails in the simulated catalogue.

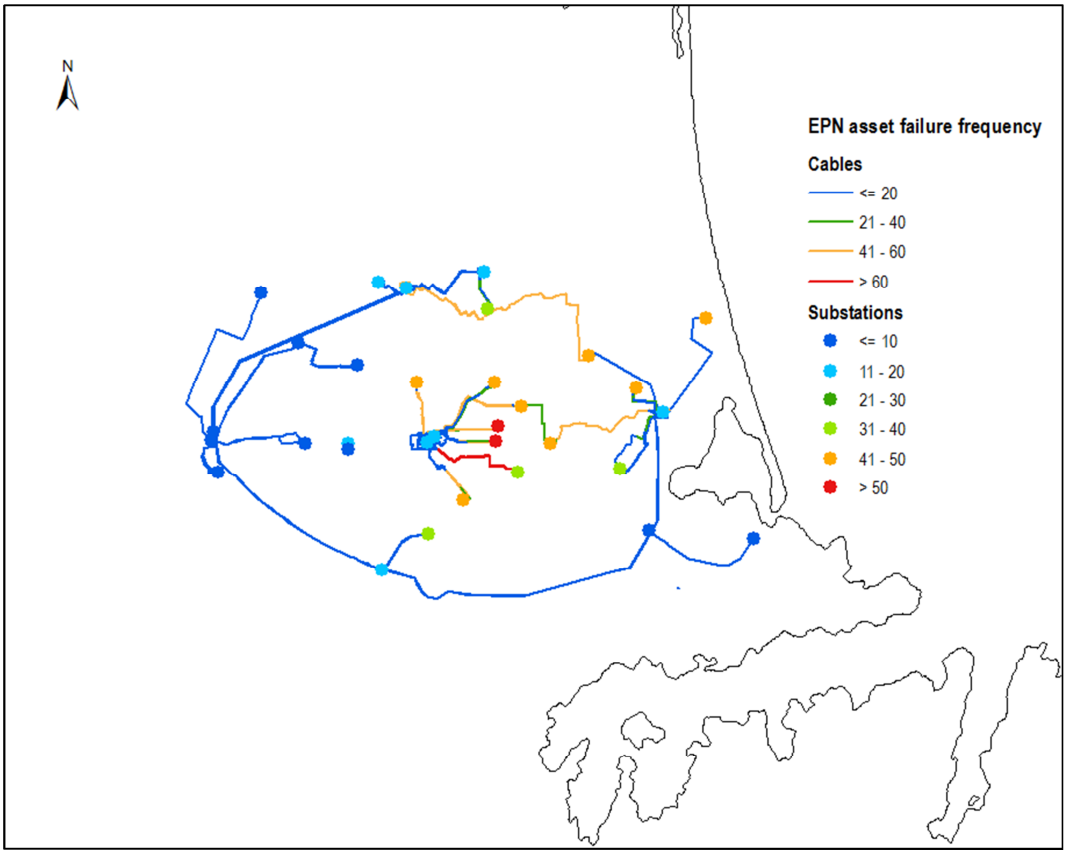

Figure 5.5 – Frequency of failures per electric power network asset in simulated catalogue of 14,428 events

Two cables fail in more than 60 events. Both of these cables run from the Addington GXP to the Milton district/zone substation and are coincident, and hence indistinguishable, in Figure 5.5. Both cables are constructed with oil-filled insulation, which relates to the 'Other' category for the repair rate model derived in section 4.3 and hence is subject to higher repair rates than other insulation materials. A further 14 cables fail in more than 40 events, including seven constructed from oil-filled insulation, five constructed with PILCA insulation and surprisingly two constructed with the more reliable XLPE insulation. It is noted however that both of these XLPE cables are located in the north eastern part of the city (one from the Bromley GXP and one from the Dallington district/zone substation), in the region most frequently affected by liquefaction, as shown in Figure 5.4. Two substations fail in more than 50 events, including the Oxford-Tuam and Montreal district/zone substations in the city centre. A further eight substations fail more than 40 times, including the Armagh, Dallington, Fendalton, Knox, Lancaster, Pages-Kearneys, Rawhiti and Spreydon district/zone substations, and it is notable

that these are mainly in the central and north eastern parts of the city where liquefaction is most frequent.

The seismic risk assessment indicates that the electric power network only suffers a loss of performance in 228 of the 14,428 events. Whilst there are an additional 37 events in which at least one asset fails, the provision of redundancy in the network means that this damage does not result in any customers being disconnected. The distribution of the loss metric for the 228 performance-affecting events is shown in Figure 5.6. In the majority of these events, the disruption to customers is relatively low (<25%), although there are a small number of events that do cause major disruption to the system. One event, the M_W 6.7 event that results in maximum damage to cables and substations, causes 100% customer disconnection. The other three events that cause over 80% customer disconnection are a M_W 7.5 event from the Awatere SW fault, and M_W 7 and M_W 6.2 earthquakes directly under Christchurch.

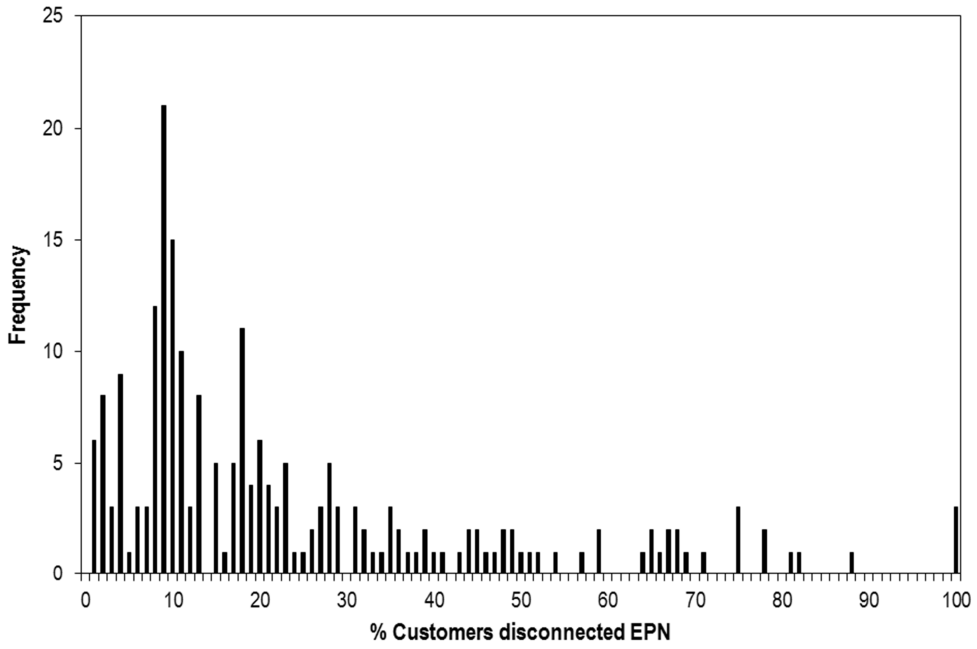

Figure 5.6 – Distribution of loss metric for 228 events that affect the performance of the electric power network

Figure 5.7 identifies how frequently each meshblock loses supply of electric power. Every meshblock becomes disconnected at least ten times and the maximum observed disconnection frequency is 124. 99 meshblocks become disconnected at least 100 times and these are

concentrated in the city centre where the substations are observed to be more frequently damaged.

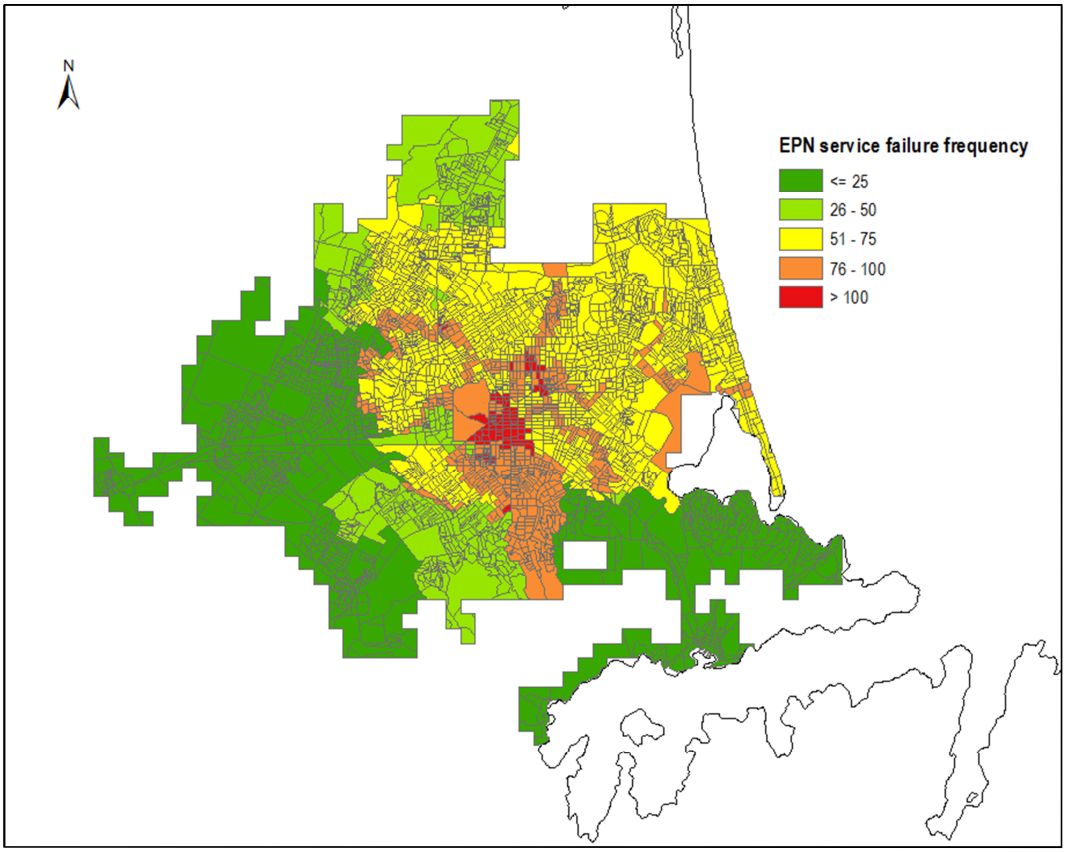

Figure 5.7 – Frequency of electric power service failures in simulated catalogue of 14,428 events by meshblock

5.1.3 Water supply system – direct physical damage

Table 5.5 and Table 5.6 summarise the effect of each event in the simulated catalogue on the water supply system assets in terms of the number of failed pumping stations (from a total of 154 including 113 primary pumping stations and 41 booster pumping station sites) and the number of failed (broken) pipes (from a total of 6,161). For the pumping stations, the numbers only include inherent failures due to direct physical damage and do not consider failures due to loss of electric power supply. Since the fragility of primary pumping station is considered to be implicit in the fragility of wells, a primary pumping station is considered to be operational if at least one of the wells from which it pumps is operational. A booster pumping station site is considered to be operational if at least one of the pumps within it is operational.

Table 5.5 – Frequency distribution of number pumping station failures per event in simulated catalogue

No. of pumping stations affected	Frequency
0	13928
1 – 10	377
11 – 20	58
21 – 30	26
31 – 40	11
41 – 50	11
51 - 60	7
61 – 70	3
71 – 80	2
81 – 90	2
> 90	3

Table 5.6 – Frequency distribution of number pipe failures per event in simulated catalogue

No. of pipes affected	Frequency
0	0
1 – 20	51
21 – 40	11469
41 – 60	2602
61 – 80	85
81 – 100	47
101 - 150	64
151 – 200	39
201 – 300	35
301 - 400	21
401 – 500	7
501 – 1000	7
> 1000	1

The results show that in the vast of the majority of events, direct physical damage does not affect the functionality of pumping stations, with failures only observed in 500 events (3.5% of all events). Conversely at least one pipe fails in every event. The largest number of pumping station failures in a single event is 129 (84%) and the largest number of pipe failures is 1,181 (19%), both caused by the same M_W 6.7 event that is seen to lead to 252 grid squares experiencing liquefaction and maximum damage to the electric power network assets.

Figure 5.8 presents the failure occurrence frequency data by asset, with the colours representing the number of times each asset fails in the simulated catalogue.

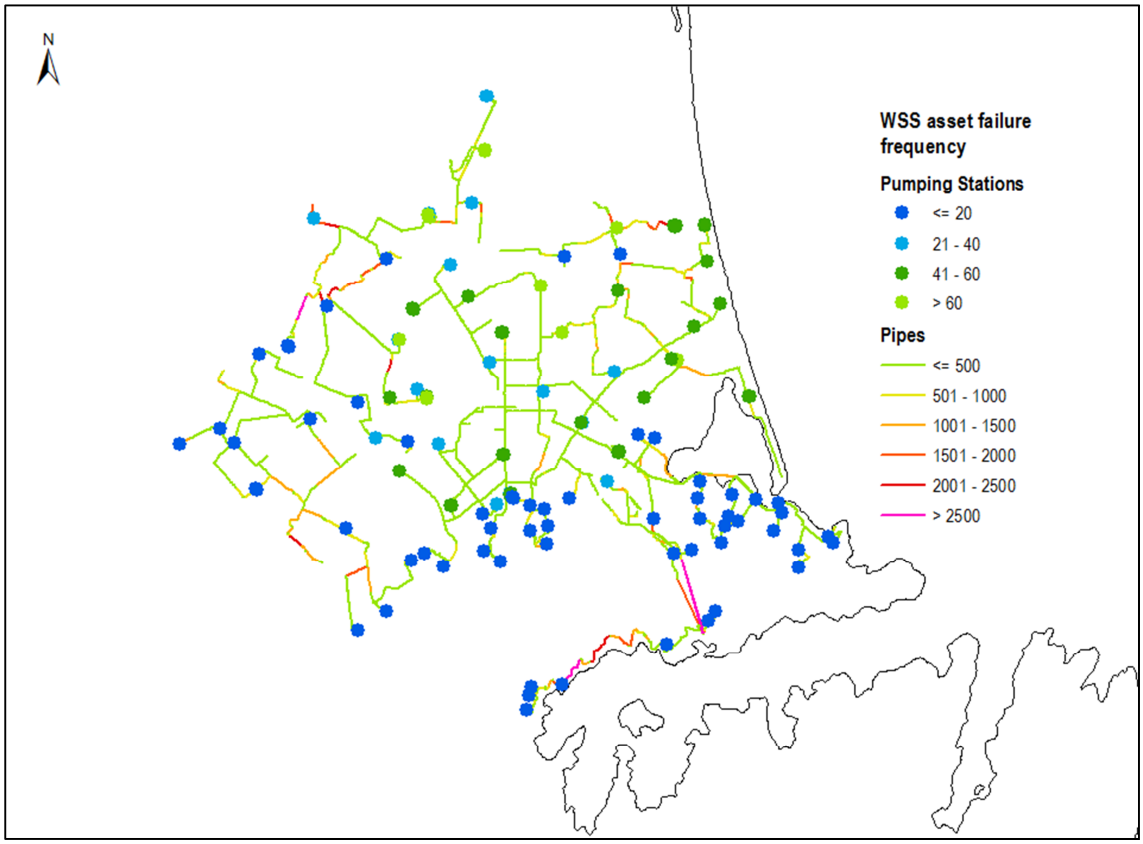

Figure 5.8 – Frequency of failures per water supply system asset in simulated catalogue of 14,428 events

Twelve pipes fail in more than 2,000 events, including four that fail more than 2,500 times (with maximum failure frequency of 2,857). These include seven pipes constructed from PVC, five pipes constructed from UPVC and two cast iron pipes. The most frequently damaged pipes are spread across the city but there is a noticeable cluster in the Lyttelton and Governor's Bay on

the southern coast and in the north western part of the city. 24 substations fail in more than 60 events, including eight that fail in more than 70 events. The maximum failure frequency is 77 occurring at the Thompsons Road 1 pumping station in the far north of the city. The other seven pumping stations with a failure frequency above 70 are Hills Road 2 and 3, Edgeware Road 1, Thompsons Road 2, Blighs Road 1 and Averill Street 3 and 4. All of these are located in the northern half of the Central pressure zone except for the Thompson Road stations, which are located in the North West pressure zone. The lower reliability of the water supply system as compared to the electric power network is confirmed by the system performance results, where every event is seen to cause some loss of performance, even when interdependencies are not taken into account. The data in Table 5.5 and Table 5.6 show that in percentage terms, a higher proportion of pumping stations suffer failure than pipes in the most damaging earthquakes. However, the median values for the fragility functions for water supply system nodes imply that in general very high accelerations are needed to cause these assets to fail. What this suggests is that whilst the fragility of pumping stations becomes a critical factor in strong high return period earthquakes, the main cause of the general vulnerability in this system across all events is likely to be the fragility of the pipes. Since they are susceptible to ground shaking as well as to liquefaction, they are inherently more vulnerable than buried cables, which partially explains why so many more events affect the water supply system than affect the electric power network. This vulnerability was observed in the Canterbury earthquake sequence and is reflected in the repair rate functions derived for this study. In terms of overall system performance, all 14,428 events in the catalogue result in some customer disconnection due to direct physical damage (the smallest value is 5%). Figure 5.9 shows the frequency distribution of the performance metric across all events. It can be seen that in the vast majority of events, the performance metric is in the range 10% to 30%. Only 210 events (1.5%) lead to more than 30% of customers being disconnected. 68 events lead to more than 50% of customers being disconnected. While only four events lead to more than 80% of customers losing electric power, this occurs nine times for the water supply system. Figure 5.10 identifies how frequently each meshblock is observed to lose water supply connection.

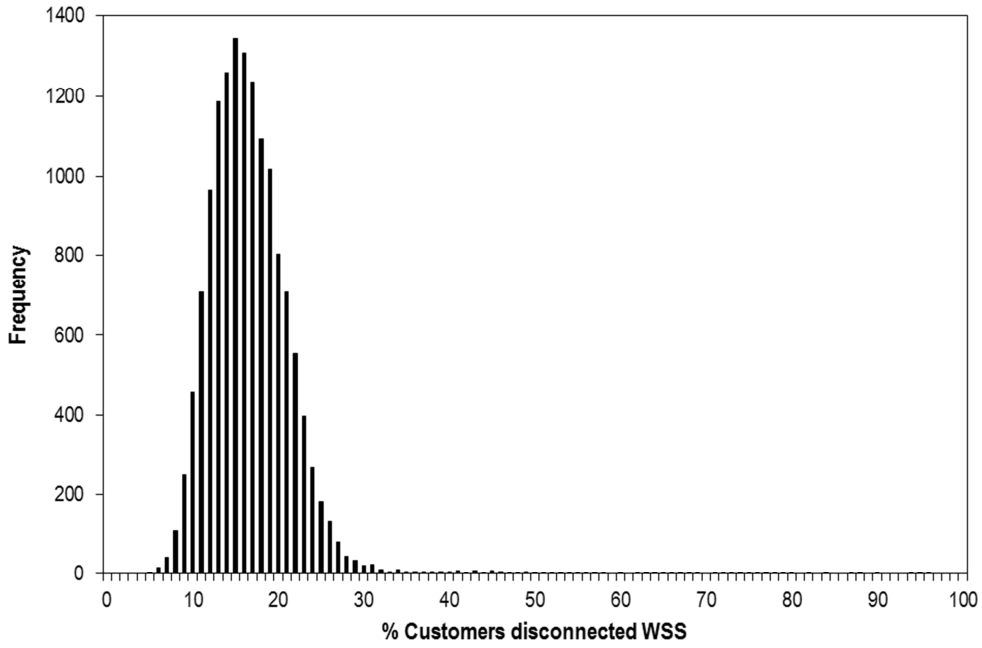

Figure 5.9 – Distribution of loss metric for 14,428 events that affect the performance of the water supply system

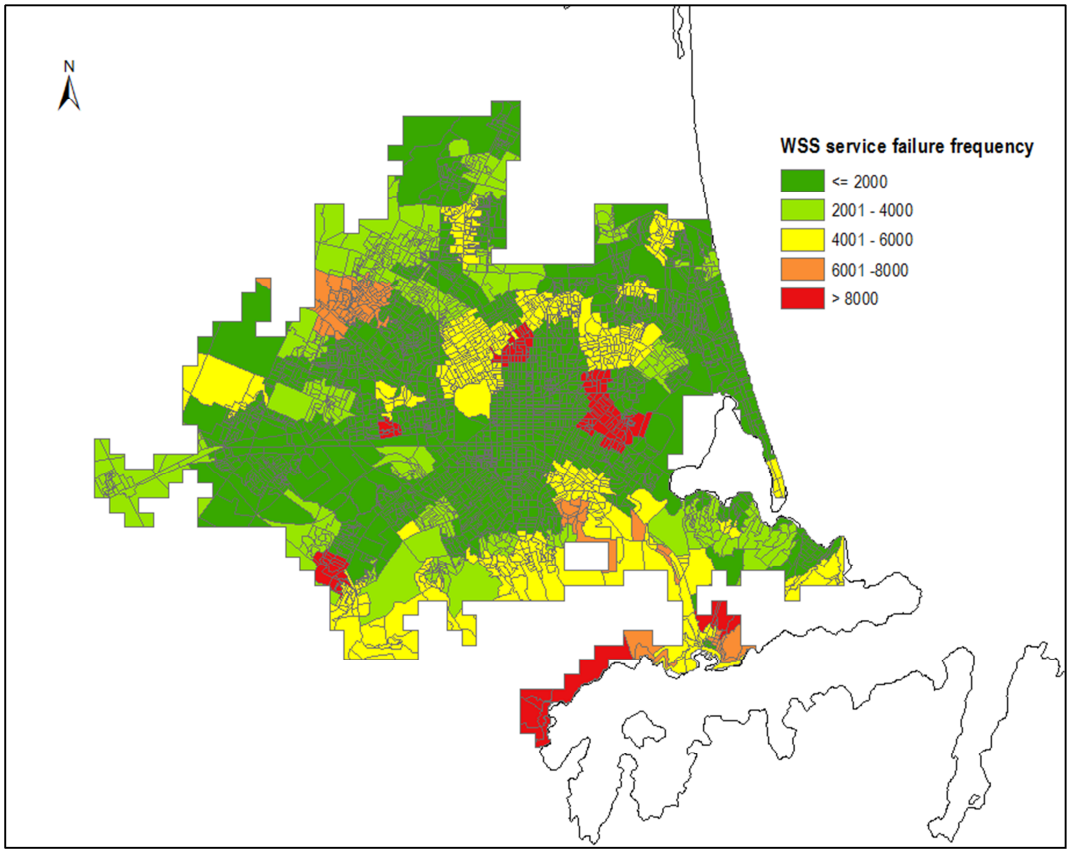

Figure 5.10 – Frequency of water supply service failures in simulated catalogue of 14,428 events by meshblock

Every meshblock experiences failure at least twice, but while there are 512 meshblocks that experience failure fewer than 100 times, there are also 142 meshblocks that suffer service failure in over 8,000 events. These occur in six clusters across the city: three locations around the city centre (Riccarton, Edgeware and Linwood), all of which are classified as having 'Moderate' or 'High' liquefaction susceptibility; Halswell in the south west of the city, which is classified as having 'Moderate' liquefaction susceptibility; and two areas on the southern shore (Lyttelton and Governor's Bay) where the ground is rocky and so there is no liquefaction risk. In these areas however, there is almost no redundancy in the system, with meshblocks connected to the water sources in the main city only by long chains of individual pipes. Therefore only one pipe in the chain needs to break for the meshblock to lose supply, and hence it is unsurprising that private water tanks are prevalent in this area.

5.1.4 Water supply system – effect of interdependency

The effect that loss of electric power has on the performance of the water supply system is accounted for by reclassifying the functionality of undamaged pumping stations that are not located in an operational power supply zone and do not have access to a back-up power supply. There are 200 events in which consideration of interdependency causes the number of pumping station failures to increase. Since the electric power network only experiences loss of performance in 228 events, it is observed that the water supply system is affected in 88% of the events in which there is potential for the interdependency to be activated. This section presents a summary of the effect of interdependencies in these events, although more detailed performance data corresponding to each event are provided in Appendix E. The largest increase in pumping station failures in a single event is 67. Amongst the 200 events in which the pumping station failure rate is affected by power loss, the mean increase is 15, which is approximately 10% of the total number of pumping stations. Table 5.7 summarises the number of pumping station failures per event when interdependency is considered. The number of events in which there are more than 20 pumping station failures increases from 65 to 123 due to interdependency and

there are now also 10 events in which more than 100 pumping stations fail. The largest number of failures in a single event is 144 (out of 154).

Table 5.7 – Revised distribution of number of pumping station failures with consideration of interdependency

No. of pumping station failures	Frequency	
	Without interdependency	With interdependency
0	13928	13921
1 – 10	377	331
11 – 20	58	53
21 – 30	26	45
31 – 40	11	22
41 – 50	11	14
51 - 60	7	8
61 – 70	3	6
71 – 80	2	8
81 – 90	2	4
91 – 100	2	6
> 100	1	10

The water supply system performance is then re-evaluated based on the new functionality classifications. The performance is observed to change in 188 events. Since the electric power network only experiences loss of performance in 228 events, it is observed that the water supply system is affected in 82% of the events in which there is potential for the interdependency to be activated. Amongst the 188 events in which the effects of interdependency are observed, the largest increase in the performance metric is 13.1%, which is equivalent to 43,075 customers, and the mean increase is 4.1%, which is equivalent to 13,383 customers. Since water supply system performance is unaffected by most events in the catalogue, a graphical representations of the new distribution of the performance metric would be indistinguishable from Figure 5.9. Instead, Table 5.8 summarises how the distribution of the water supply system performance

metric changes due to interdependency when grouped into bins. This shows that consideration of interdependency leads to 16 additional events in which the level of customer disconnection is in excess of 50%.

Table 5.8 – Effect of electric power network interdependency on distribution of performance metric for water supply system

% Customers disconnected	Frequency	
	Without interdependency	With interdependency
< 10	877	877
10 – 20	10916	10913
20 – 30	2425	2425
30 – 40	90	84
40 – 50	52	45
50 - 60	25	32
60 – 70	15	13
70 – 80	19	18
80 – 90	6	14
> 90	3	7

The consideration of interdependency increases the frequency with which 1,601 (48%) of the meshblocks become disconnected from the water supply system. The largest observed increase in meshblock failure frequency is 68 events. Figure 5.11 shows the change in frequency in each meshblock. It is notable that most of the affected meshblocks are towards the edges of the city where there is a lower level of redundancy in the electric power network. Across all meshblocks, the mean increase is 8 events, while the mean increase amongst the 1,601 affected meshblocks is 16 events. In percentage terms, the largest observed increase is 500%. Table 5.9 summarises the distribution of the absolute and percentage changes in failure frequency amongst the affected meshblocks.

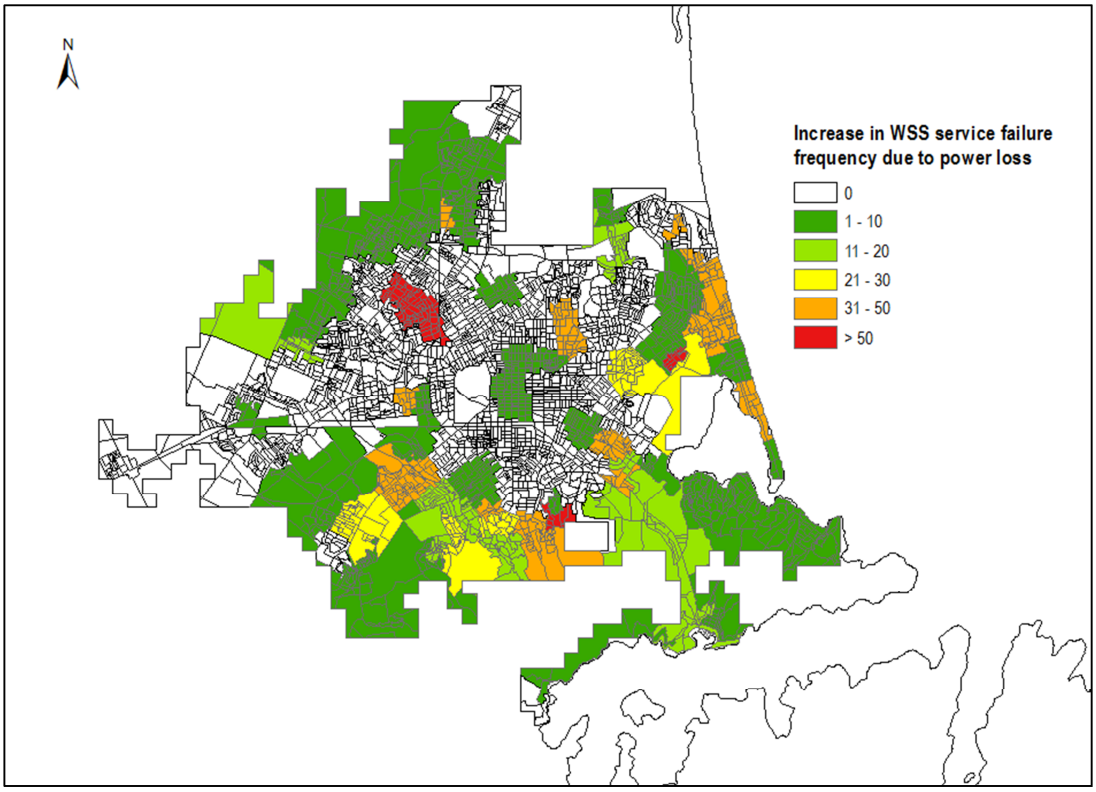

Figure 5.11 – Increase in frequency of meshblock water supply service failures in simulated catalogue due to interdependency with electric power network

Table 5.9 – Summary of changes in meshblock failure frequency caused by interdependency

Absolute increase in failure frequency (no. of events)	No. of meshblocks	Relative increase in failure frequency (%)	No. of meshblocks
1 – 10	959	0 – 10	1238
11 – 20	169	10 – 20	112
21 – 30	104	20 – 30	2
31 – 40	202	30 – 40	14
41 – 50	78	40 – 50	71
51 - 60	71	50 – 60	12
61 – 70	18	60 – 70	21
71 – 80	0	70 – 80	71
81 – 90	0	80 – 90	0
91 – 100	0	90 – 100	0
> 100	0	> 100	60

5.1.5 Risk metrics

In engineering studies, results from seismic risk assessments of critical infrastructure systems are commonly presented as system risk curves. These plot annual exceedance probabilities against system performance. The curves for the electric power and water supply systems are shown in Figure 5.12, where the system performance metric has been annualised by taking the maximum value observed to occur in each year. For the water supply system, this includes curves for independent performance and both methods for interdependent performance: by electric power supply zone as part of the final developed model and by geographical proximity as a sensitivity test. The different approaches to modelling interdependency only yield different results for low frequency events and so for clarity, the annual exceedance probability is plotted on a logarithmic scale.

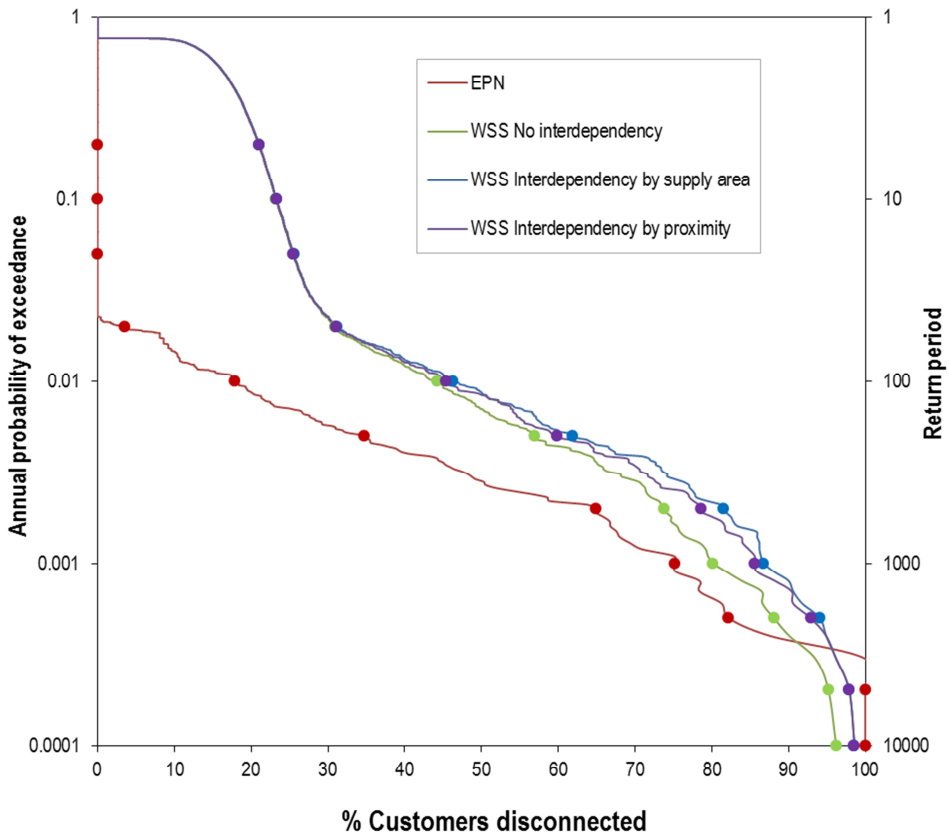

Figure 5.12 – System risk curves Christchurch electric power and water supply systems

In the insurance sector it is more common to present the loss metrics corresponding to specified return periods. The points marked on each curve in Figure 5.12 represent the performances corresponding to the following return periods, measured in years: 5, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000 and 10,000. The values corresponding to these points are listed in Table 5.10.

Table 5.10 – Performance metrics for each system corresponding to key return periods (the change in water supply system performance metric due to interdependency is indicated in brackets)

Return period (yrs)	Performance metric		
	Electric power network	Water supply system – without interdependency	Water supply system – with interdependency
10,000	100	96.2	98.6 (+2.4)
5,000	100	95.2	97.9 (+2.7)
2,000	82.2	88.1	94.1 (+6.0)
1,000	75.2	80.2	86.8 (+6.6)
500	64.9	73.8	81.6 (+7.8)
200	34.7	56.9	61.9 (+5.0)
100	17.9	44.3	46.3 (+2.0)
50	3.6	30.9	31.2 (+0.3)
20	0	25.5	25.6 (+0.1)
10	0	23.2	23.3 (+0.1)
5	0	21.0	21.0 (-)

The risk metrics highlight that whilst the electric power network is very robust with respect to weak high frequency events, significant disruption is expected for stronger low frequency events. In particular, whilst no events are predicted to cause total failure of the water supply system, for the electric power network, total failure is predicted to occur three times. For the water supply system, at moderate to high frequencies (annual exceedance probability ≥ 0.01), there is negligible difference in performance regardless of whether or not interdependencies are taken into account or how. This is because at these frequencies, there is little or no loss of

performance in the electric power network and so interdependencies are not activated. As annual exceedance probability decreases, the difference between independent and interdependent performances also increases. The difference between the two interdependency modelling methods appears relatively small. However, when compared to independent system analysis, modelling interdependency by electric power supply zones increases the performance metric by 7.7% (in relative terms) when averaged across the 228 events where the electric power network suffered loss of performance, whilst when interdependency is modelled by proximity, the increase is 11.8%, confirming that the geographical proximity method is conservative in this case study as it does not account for redundancy.

5.1.6 Mitigation strategies

An important function of a seismic risk assessment model for critical infrastructure systems is the ability to investigate the impact of potential mitigation works that could improve system reliability. This can be done by comparing the seismic risk assessment results obtained with the current system to new sets of results obtained from corresponding assessments that include proposed changes. Multiple proposed strategies can be assessed to compare effects and identify priorities.

Mitigation in the electric power network is investigated by comparing the effects of strengthening different substations. Ten potential mitigation strategies are proposed, each involving the strengthening of one substation. The ten substations with the highest failure frequencies from the original seismic risk assessment are selected as the candidates for strengthening and these are listed in Table 5.11. Each mitigation strategy test uses the same events and asset damage states as the original seismic risk assessment, with the exception of the candidate substation, which is manually set to 'No damage' for all events in order to mimic the effect of being strengthened. The system performance is then re-calculated for each event based on the revised damage states. For the electric power network, there are 228 events in which loss of performance is observed. For each mitigation strategy test, Table 5.11 summarises the mean

increase in customer connections across the 228 events and the number of events in which the mitigation strategy significantly improves system performance.

Table 5.11 – Effect on seismic risk assessment results due to strengthening of least reliable substations

Substation	Failure frequency in original seismic risk assessment	Effect of strengthening		
		Mean increase in no. of EPN customer connections	No. of events with EPN performance improvement > 5% of population	No. of events with EPN performance improvement > 10% of population
Oxford-Tuam	58	199	0	0
Montreal	54	205	0	0
Dallington	50	4784	42	1
Pages-Kearneys	50	3755	31	4
Rawhiti	49	4375	39	0
Armagh	46	683	1	0
Spreydon	46	1075	0	0
Lancaster	44	1336	0	0
Fendalton	43	2650	21	0
Knox	41	475	0	0

The results show that although the Oxford-Tuam and Montreal substations are the least reliable in terms of physical damage, its strengthening would lead to minimal improvement in system performance. This may be because they are centrally located in an area where there is greater redundancy in the network and so although they are failing frequently in the original seismic risk assessment, their failure does not cause significant loss of performance. However, the Dallington, Pages-Kearneys and Rawhiti substations, which are all located in suburban Christchurch, do all have a strong positive effect on system performance if strengthened. The average improvement corresponding to these substations is between 1% and 2% of the total population, but they also result in many individual events in which the improvement is greater than 5% and strengthening of the Pages-Kearneys substation in particular results in four events

in which the improvement is greater than 10%. It is clear that these three substations should be the priority for seismic upgrading.

In the water supply system, mitigation is investigated by comparing the effects of strengthening the ten least reliable pumping stations as listed in Table 5.12. The method is the same as for the electric power network, with the damage state of the candidate pumping stations manually set to 'No damage' for all events. To reduce the computation time, and because the impact of mitigation is most important in highly damaging earthquakes, the mitigation strategies are only tested for events in the original seismic risk assessment where the loss of performance in the water supply system is observed to be at least 25% after interdependencies are modelled. This results in a total of 617 events. For each mitigation strategy test, Table 5.12 summarises the mean increase in customer connections across the 617 events and the number of events in which the mitigation strategy significantly improves system performance.

Table 5.12 – Effect on seismic risk assessment results due to strengthening of least reliable pumping stations

Pumping station	**Failure frequency in original seismic risk assessment**	**Effect of strengthening**		
		Mean increase in no. of WSS customer connections	**No. of events with WSS performance improvement > 5% of population**	**No. of events with WSS performance improvement > 10% of population**
Thompsons Road 1	77	71	0	0
Edgeware Road 1	74	41	0	0
Blighs Road 1	74	41	0	0
Hills Road 2	73	41	0	0
Hills Road 3	73	41	0	0
Thompsons Road 2	73	71	0	0
Averill Street 4	72	51	0	0
Averill Street 3	71	51	0	0
Woodham Road	70	64	0	0
Picton Avenue 2	69	104	0	0

The results show that for all of the pumping stations the average improvement in system performance from strengthening is very small. In terms of individual events, the maximum observed improvement is just 6,950 customers (just over 2%) at Woodham Road. This reinforces the conclusion from the results presented in Figure 5.8 that it is damage to pipes rather than pumping stations that is driving loss of performance in the water supply system and therefore pipes should be prioritised in any seismic mitigation strategy.

5.2 Interdependency quantification

This section presents the analysis of the seismic risk assessment results to quantify the interdependency between the electric power network and water supply system. The sections begins with the application of the proposed 'linear slope method', as described in section 3.3, before carrying out sensitivity tests on the observed interdependency response by varying the fragility of assets and the availability of back-up power in the water supply system. The section then presents the results from the application of the simplified 'extreme event method' for quantifying interdependency response, which is also proposed and described in section 3.3.

5.2.1 Linear slope method

As described in section 3.3, the proposed interdependency index, $I_{WSS|EPN}$ can be estimated from the slope of the line of best fit on a plot of the interdependency response in the water supply system against electric power network performance. Interdependency only becomes a consideration in cases where there is some loss of performance in the electric power network, and so Figure 5.13 plots the interdependency response in the water supply system against electric power network performance for the 228 events in which this is the case (for other events the points would all plot at the origin).

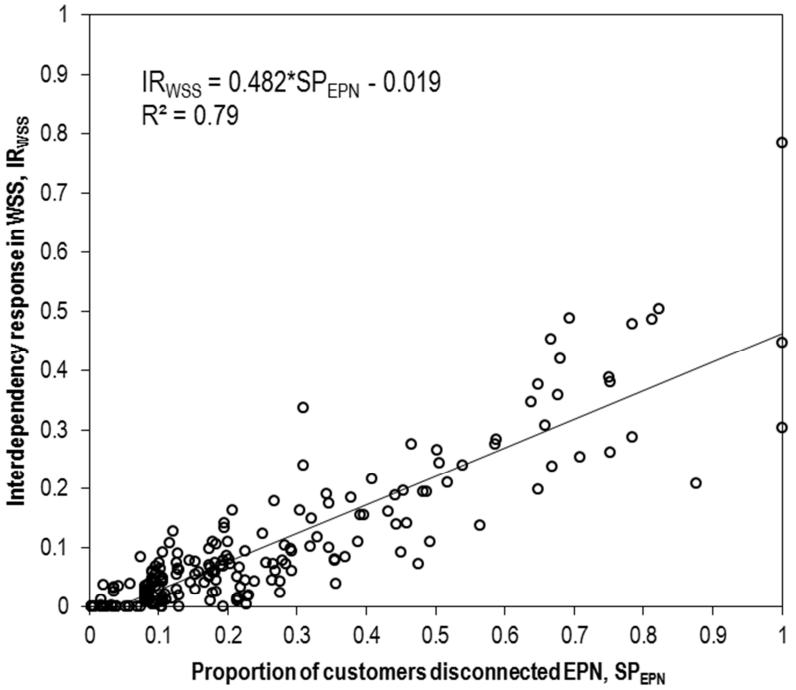

Figure 5.13 – Plot of observed interdependency response against electric power network performance for 228 events in seismic risk assessment catalogue where electric power network suffered loss of performance

In 166 events the interdependency index is less than 0.1 and in a further 32 events it is less than 0.2. The index exceeds 0.5 in only two of the events, which indicates that in general the impact of electric power network performance on the water supply system is low. The plot shows the general trend that the interdependency response increases as the performance of the electric power network worsens. Ordinary least squares linear regression is applied to derive a predictive function for interdependency response based on electric power network performance as printed on the graph. Although there is some scatter around the best fit line, the regression is significant at the 5% level and the calculated coefficient of determination (R^2) indicates that the electric power network performance explains approximately 80% of the variance in the interdependency response in the water supply system and confirms that a linear function is a good fit between the datasets. Therefore, using the slope method, the seismic risk assessment yields a mean value for the proposed interdependency index of 0.48, with a 95% confidence interval between 0.45 and 0.51. However, a sample size of 228 may be considered to be relatively small and inconclusive. This is particularly true here, since the majority of

observations are clustered in the bottom left corner and so the predictive power of the model is effectively weighted towards low damage scenarios, at the expense of the more devastating scenarios for which it could be argued that the prediction of interdependency response is more important. Therefore, further interdependency analysis is undertaken to attempt to replicate the result for the interdependency index using different event sets.

The proposed model for predicting interdependency response (the adapted Leontief input-output method) is deterministic, i.e. the same value of interdependency response will always be predicted for a given value of electric power network performance. Therefore, the process for deriving the predictive function does not need to be based on a stochastically representative event catalogue. The only requirement is that a spread of electric power network performances are included as this is the only independent variable. In fact it could be argued that a non-stochastic catalogue is better for deriving the interdependency response model, as by ignoring event probability, it is likely to include more high magnitude events and likely to provide greater balance between predictive power for high frequency low impact events and predictive power for low frequency high impact events. Here, the interdependency quantification analysis is therefore repeated for four new catalogues, each containing 250 events generated non-stochastically from the Christchurch earthquake source model. Each event is subject to the full seismic risk assessment methodology incorporating hazard, fragility and system performance analyses and nodal interdependency simulation. The process for generating each catalogue is iterative, as events that lead to no loss of performance in the electric power network are discarded and replaced until all 250 events contribute to the interdependency analysis. To accelerate the process and increase the likelihood of generating events that cause loss of performance to the electric power network, the source model is simplified to include only events of $M_W \geq 6$ within 50 km of Christchurch. Figure 5.14 shows the geographical distribution of events in the four catalogues and the frequency distribution of magnitudes is summarised in Figure 5.15.

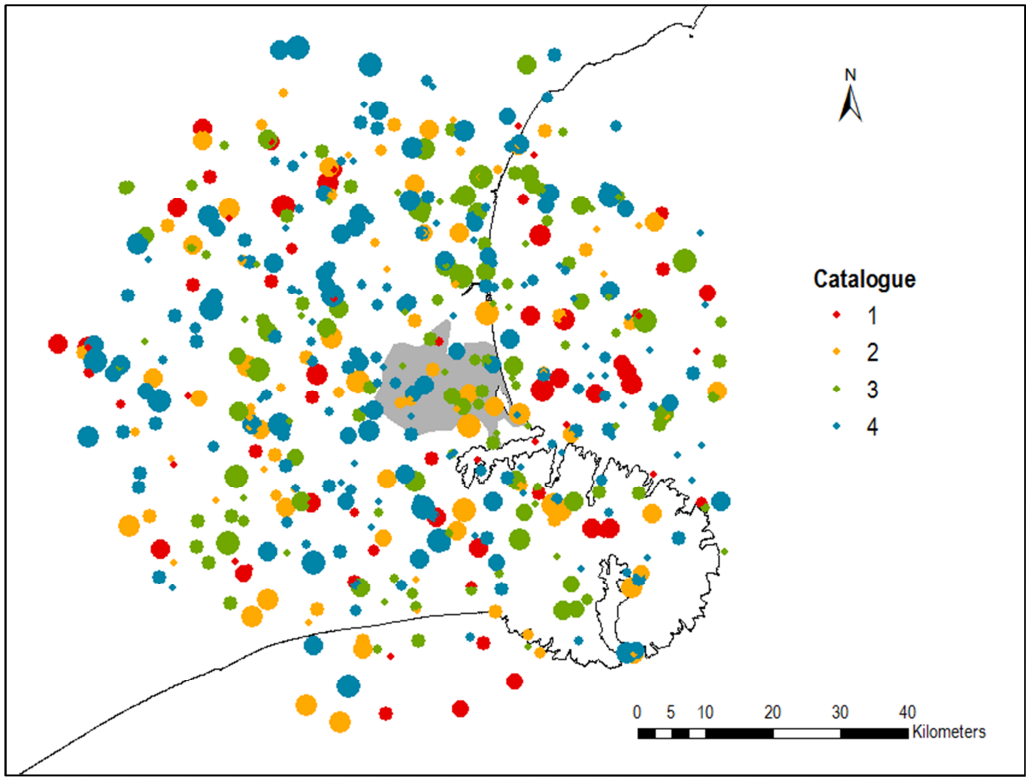

Figure 5.14 – Geographical distribution of events in four non-stochastic catalogues for interdependency analysis (the size of the circles increase with moment magnitude ranging from 6.0 to 7.3)

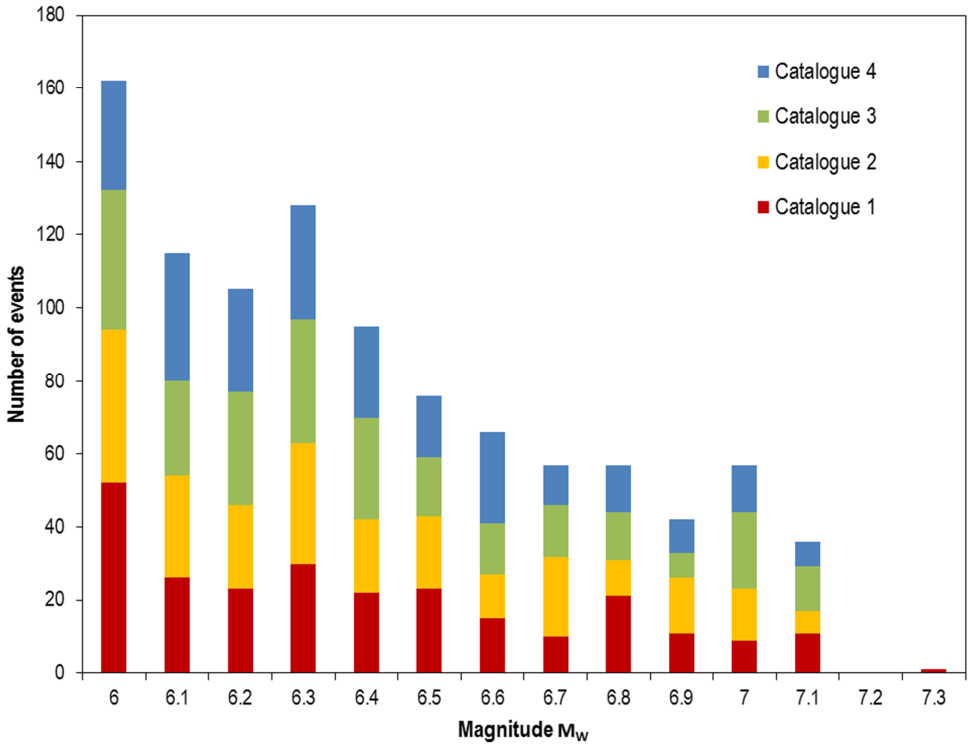

Figure 5.15 – Frequency distribution of earthquake magnitudes in four non-stochastic catalogues for interdependency analysis

Figure 5.16 shows the distribution of system performances resulting from the events generated for the four new catalogues combined. As expected, given the high reliability of the electric power network, there is a bias in the catalogue towards events with good system performances. However, there is a greater spread of performances compared to the 228 events from the seismic risk assessment catalogue in which the electric power network suffered loss of performance (see Figure 5.6). Of those 228 events, 62% have a performance metric less than or equal to 0.2, compared to 39% in the new catalogues. Just 12% of events have a performance metric greater than 0.5 in the seismic risk assessment catalogue compared to 30% in the new catalogues. This improved spread suggests that the new catalogues may produce models that are more robust for higher electric power network performance metrics. For the water supply system, significantly more than half the events result in performances worse than 0.5. There are no events with performance lower than 0.1, but this is a consequence of the low reliability of the system, combined with the need for events that cause performance loss in the more reliable electric power network.

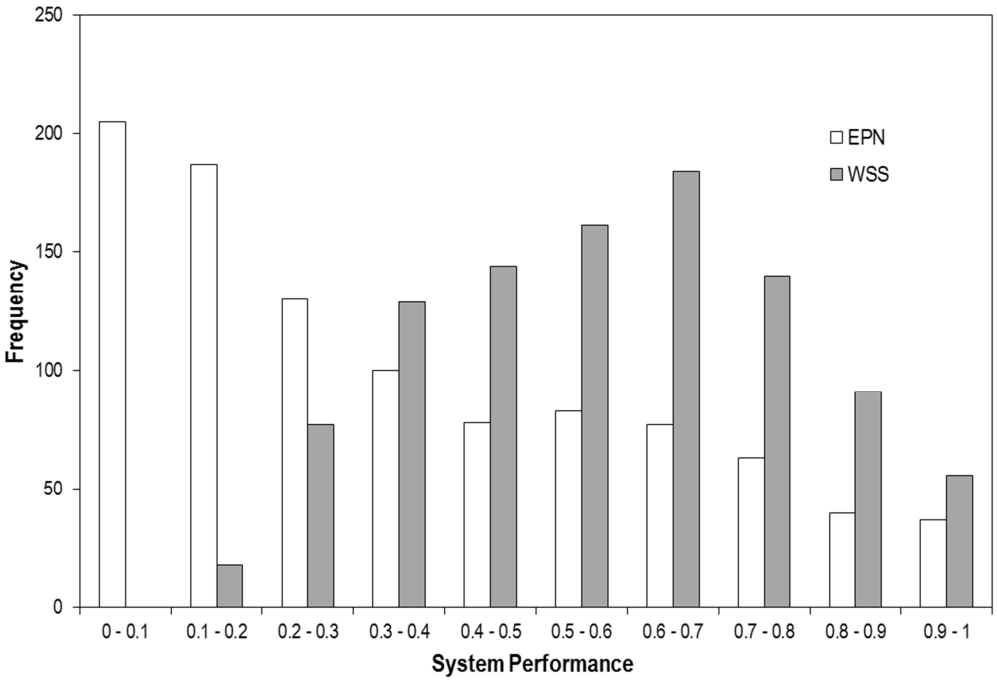

Figure 5.16 – Distribution of system performances resulting from events in supplementary catalogue for interdependency analysis

Figure 5.17 plots the interdependency response in the water supply system against electric power network performance for the four new catalogues.

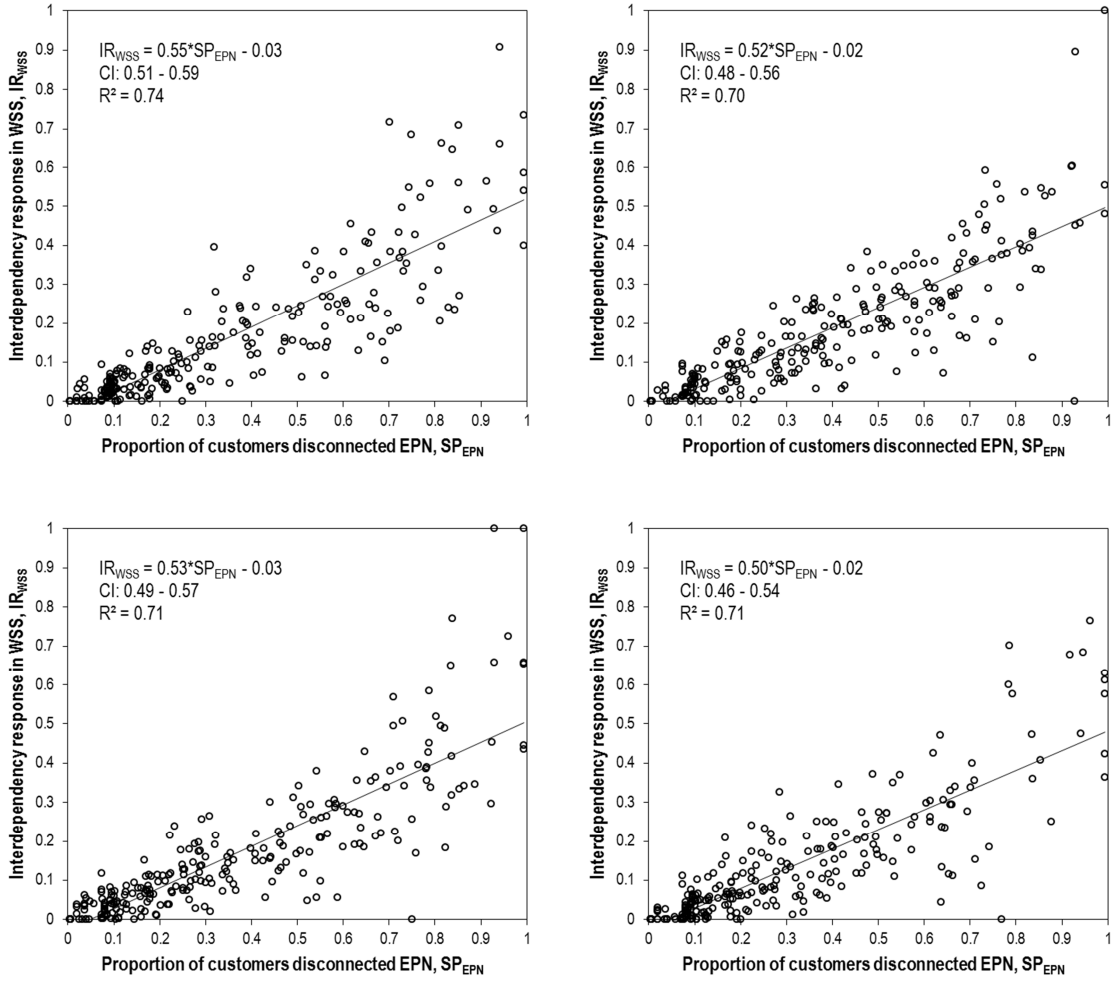

Figure 5.17 – Plot of observed interdependency response against electric power network performance for four non-stochastic 250 event catalogues

In all the catalogues the interdependency response is less than 0.5 for the majority of events with only 49 instances exceeding this value (combined across all catalogues). There are a lower proportion of events with the lowest values of interdependency response (<0.1), compared to the seismic risk assessment catalogue, but generally all four non-stochastic catalogues indicate that the impact of the electric power network on the water supply system is only moderate. New linear predictive functions are printed on each graph, including the 95% confidence interval for the slope. There is more scatter in these datasets, as reflected by the lower coefficients of

determination. However, they are all greater than 0.7 and so still indicative of a good linear fit and all the regressions are significant at the 5% level.

The interdependency indices derived from these catalogues range from 0.5 to 0.55, which indicates that there is some consistency in the value of the derived index, although it is slightly higher than the value derived from the seismic risk assessment catalogue. This is an expected consequence of the new catalogues having a higher proportion of high impact events but the significance of this difference needs to be ascertained. The first step is to determine whether the four slopes derived from non-stochastic catalogues are significantly different from each other by comparing the two datasets with the largest (top left) and smallest (bottom right) slopes. Since the sample size is large (>30), a two-sample Z-test is applicable. The null hypothesis is that the slopes from these two datasets are the same and the Z-statistic is given by Eq. 5.1, where β_{IR1} and β_{IR2} are the slopes of the models (0.55 and 0.50), and SE is the standard error of the slopes (0.020 and 0.021).

$$Z = \frac{\beta_{IR1} - \beta_{IR2}}{\sqrt{SE^2_{\beta_{IR1}} + SE^2_{\beta_{IR2}}}} \tag{5.1}$$

The calculated Z-statistic for the two models is 1.59. At the 5% significance level, the critical value of Z is 1.96, giving a p-value of 0.11. The null hypothesis cannot therefore be rejected and so it is assumed that all four slopes from the non-stochastic catalogues are not significantly different from each other. This means that the four catalogues can be combined to derive a more robust model from a larger sample of 1,000 events. The combined scatter plot is shown in Figure 5.18, with the new best linear function fit. It is notable that the confidence interval for the interdependency index does not overlap with the value derived from the stochastic catalogue. Since the sample sizes are large, a two-sample Z-test is applied to formally test whether the difference between the stochastic and non-stochastic catalogues is significant, with a null hypothesis that they are not significantly different.

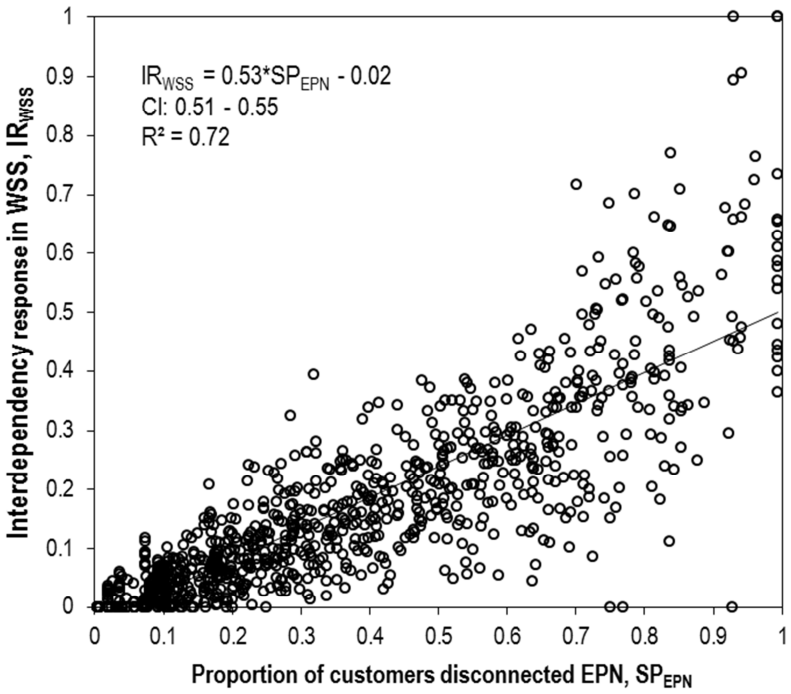

Figure 5.18 – Plot of observed interdependency response against electric power network performance for combined non-stochastic catalogue of 1,000 events

For the stochastic model, the slope is 0.48 and the standard error of the slope is 0.016. For the non-stochastic model, the slope is 0.53 and the standard error of the slope is 0.010.The calculated Z-statistic for the two models is 2.33. At the 5% significance level, the critical value of Z is 1.96, giving a p-value of 0.02. Therefore the null hypothesis is rejected and the interdependency indices derived from the stochastic and non-stochastic catalogues are significantly different. Given the deterministic nature of the interdependency index and the fact that the spread of events is wider in the non-stochastic catalogue, this should then be the preferred method of calculation.

5.2.2 Non-linear interpretations

The analysis so far has assumed that the interdependency response function is linear, since this is a basic premise of input-output interdependency modelling. However, visual inspection of the plots in Figure 5.17 and Figure 5.18 show a pronounced upwards kink when the performance metric in the electric power network exceeds 0.8, suggesting that an exponential or polynomial

curve may provide a better fit than a linear function. In general a simple exponential relationship can be fit to a dataset by applying a logarithmic transformation to both the dependent and independent variables and then carrying out linear regression on the transformed data. However, since observations of the interdependency response include zero values, this method cannot be applied. Polynomial curves do not require transformation for fitting and do accommodate zero values in both variables, and so three forms of polynomial curve are proposed for fitting to the non-stochastic catalogue results as shown in Eq.'s 5.2 to 5.4, where IR_{WSS} is the interdependency response in the water supply system, SP_{EPN} is the performance metric for the electric power network and a_{IR}, b_{IR} and c_{IR} are coefficients to be determined.

Polynomial model 1: $$IR_{WSS} = aSP_{EPN}^2 + b \tag{5.2}$$

Polynomial model 2: $$IR_{WSS} = aSP_{EPN}^2 + bSP_{EPN} + c \tag{5.3}$$

Polynomial model 3: $$IR_{WSS} = aSP_{EPN}^b + c \tag{5.4}$$

Polynomial regression is a special case of multiple linear regression and as such can be analysed using the ordinary least squares method. The best fits for each of the three polynomial forms is shown in Figure 5.19 alongside the linear function for the non-stochastic catalogue, although the best fits for polynomial models 2 and 3 are almost indistinguishable. All regressions are significant at the 5% level and the coefficients of determination are all above 0.7. The Akaike Information Criterion (AIC) (Akaike, 1974) is a tool used for model selection where a variety of model formulations have been fit to the same dataset. Assuming that there exists a 'true' model for predicting interdependency response, then AIC measures the relative quality of each of the candidate models to select the model that minimises loss of information relative to the 'true' model. For each model, the value of AIC can be calculated from Eq. 5.5, where n_{IR} is the size of the dataset, k is the number of independent variables in the model, and SS_{res} is the residual sum of squares for the model as determined in the regression. The best candidate model is the model with the lowest value of AIC.

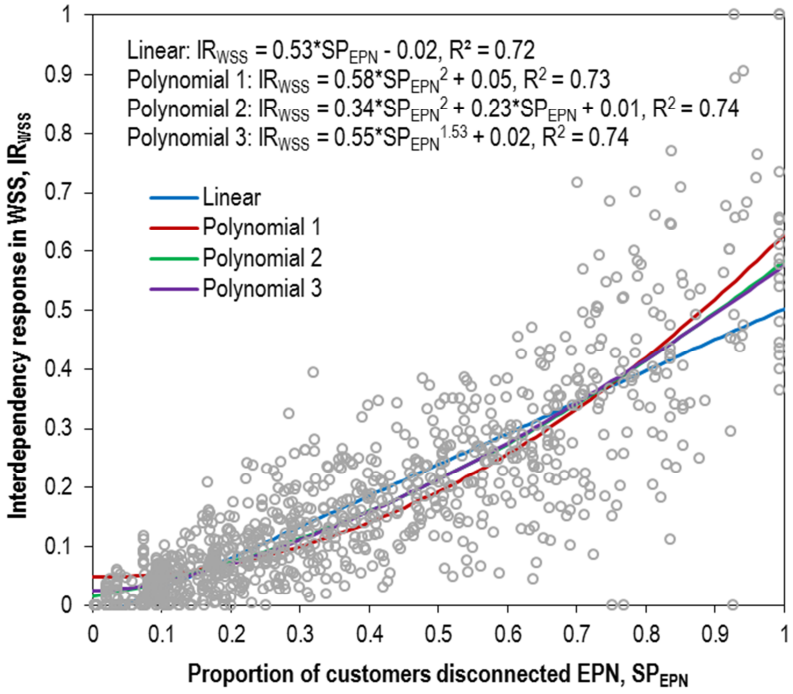

Figure 5.19 – Plot of observed interdependency response against electric power network performance for combined non-stochastic catalogue of 1,000 events with best fit linear and polynomial models

$$AIC = n_{IR} \ln\left(\frac{SS_{res}}{n_{IR}}\right) + 2\left(k+1\right) \tag{5.5}$$

The relative quality of each model can be determined by estimating the likelihood of each model by Eq. 5.6, where L is the likelihood, which is proportional to the term on the right hand side and where ΔAIC is the difference between the AIC value for the model and the minimum AIC value across all models.

$$L \propto e^{-0.5\Delta AIC} \tag{5.6}$$

The best model by definition has a likelihood of 1. The relative weight of each model is the likelihood of the model expressed as a proportion of the sum of all model likelihoods. The results of the AIC analysis are shown in Table 5.13, which shows that polynomial model 2 provides the best representation of the data.

Table 5.13 – Results of model selection by Akaike Information Criterion

Model	AIC	Likelihood $\propto$	Weight
Linear	-4863	< 0.01	< 0.01
Polynomial 1	-4894	< 0.01	< 0.01
Polynomial 2	-4931	1	0.96
Polynomial 3	-4924	0.04	0.04

In addition to AIC analysis, an ANOVA (analysis of variance) test between the linear model and polynomial model 2 is applied to test whether the difference between the linear model and polynomial model 2 are significant. The test statistic is the residual sum of squares of each model and is assumed to have an F-distribution. The null hypothesis is that the residual sum of squares are not significantly different, but the ANOVA test yields a p-value < 0.01, so the null hypothesis is rejected and the difference in the two models is confirmed to be significant. However, there are practical considerations that must also be considered when assessing the applicability of a model that departs from the basic linear function assumed by Haimes and Jiang (2001).

The development of the interdependency response function, and by extension, the proposed interdependency index serves two purposes. The first purpose is that a function for interdependency response can be used as a simplified method for simulating interdependencies in future risk analyses where detailed nodal modelling of interactions may not be possible (e.g. lack of data) or undesirable (e.g. cost of resources outweigh perceived benefits of greater precision). If interdependencies only need to be simulated for one pair of systems with a one-way relationship, then it is relatively straightforward to apply response functions with more complex forms. For example, to simulate interdependencies with the linear model found in Figure 5.19, the equation is subsumed into Eq. 3.106 to give Eq. 5.7. To apply polynomial model 2 instead, it is simply a case of substituting the polynomial function for the linear equation, as shown in Eq. 5.8. It is noted that in both cases the intercept term is ignored to

ensure that there is no interdependency response when there is no loss of performance in the electric power network.

$$SP_{WSS} = 0.53SP_{EPN} + SP'_{WSS} \tag{5.7}$$

$$SP_{WSS} = \left(0.34SP_{EPN}^{2} + 0.23SP_{EPN}\right) + SP'_{WSS} \tag{5.8}$$

The situation becomes more complicated if there are multiple systems or if there are feedback loops. To find the final performances manually requires iterating through each pair of systems individually and repeatedly until a steady state is achieved across all systems. As discussed in section 2.6.2, the problem is simplified by combining all the relationships into matrix form and solving for a vector of the final system performances. Referring back to Eq. 2.47, the applicability of the matrix method requires the unknown terms inside the final system performance vector (**x**) to be the same as the predictor variables in the interdependency response function. For example, if Eq. 5.8 is written in matrix form, it would become Eq. 5.9 and the solution for the final system performances would be given by Eq. 5.10.

$$\begin{pmatrix} SP_{EPN}^{2} \\ SP_{EPN} \\ SP_{WSS} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.34 & 0.23 & 0 \end{pmatrix} \begin{pmatrix} SP_{EPN}^{2} \\ SP_{EPN} \\ SP_{WSS} \end{pmatrix} + \begin{pmatrix} SP'^{2}_{EPN} \\ SP'_{EPN} \\ SP'_{WSS} \end{pmatrix} \tag{5.9}$$

$$\begin{pmatrix} SP_{EPN}^{2} \\ SP_{EPN} \\ SP_{WSS} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -0.34 & -0.23 & 1 \end{pmatrix}^{-1} \begin{pmatrix} SP'^{2}_{EPN} \\ SP'_{EPN} \\ SP'_{WSS} \end{pmatrix} \tag{5.10}$$

These demonstrate that if the interdependency response function contains terms where the system performance is raised to some exponent other than 1, then the final system performance must also contain the same terms. In a scenario such as this, where the performance of the electric power network is included twice (to the powers of one and two), this leads to a situation where two of the unknowns are not independent of each other. This, in turn, may prevent the matrix formulation from being solvable. In this simple example with only one pair of systems

and a one-way dependency, this does not pose a problem but it may do in more complex scenarios. This could be avoided by limiting polynomial functions to a single variable term as in polynomial models 1 and 3 (Eq.'s 5.2 and 5.4). However, in a scenario where multiple systems are being analysed and there is a system (e.g. the electric power network) that supports more than one of the other systems (e.g. water supply system and natural gas system), then the interdependency response function for each relationship must be of the same form, (i.e. the performance of the electric power network must be raised to the same exponent in each function). The likelihood that the best fit model for each relationship requires the performance of the electric power network to be raised to the same exponent is low.

From a purely scientific perspective it is clear that the interdependency response functions for each pair of infrastructure systems should be described by the model that objectively best fits the data, whether it is linear, polynomial or some other from. However, other than for the most straightforward scenarios, this approach can lead to significant complexity in the application of the matrix methods to solve for the final system performances and at worst, preclude it altogether. In the latter case the analysis must then be solved manually, which can itself be very complex and resource-intensive. Whilst this may be feasible for academic research, in the commercial and public sectors, such a level of complexity may act as a barrier to adoption of these methods. Furthermore, the assumption that all models should have the same basic form may be perceived as beneficial for non-academic studies, since it allows standardisation of model derivation and application procedures for all pairs of systems.

The second purpose for developing the interdependency response function is that the slope of the best linear fit – the interdependency index – can be used as an objective measure for quantifying the interdependency between systems for risk information and management. This is useful since it allows relationships between different pairs of systems to be compared. In the linear model, and polynomial models 1 and 3, in Figure 5.19, the slope is approximately equivalent to the expected interdependency when there is total loss of electric power network performance, and so these models have the same basic meaning. However, the linear function in

Figure 5.19 underestimates the interdependency index relative to the polynomial models and ANOVA testing shows that the differences are significant. Compared to a polynomial function, a linear function is likely to always underestimate the maximum y-value, which is what the index essentially is. So although they are, by definition, measuring the same phenomenon, it is questionable whether they are directly comparable. In polynomial model 2, where there is more than one variable term in the function, neither coefficient has a practical real-world meaning. They are neither comparable to coefficients from other model forms nor meaningful in themselves. Direct comparison between different pairs of systems is therefore only valid if the metrics have the same meaning and have been derived using the same form with the system performance raised to the same exponent. Whilst this can be done with polynomial functions, it is not necessarily the case that all interdependency relationships will be describable by a polynomial function, and even if they are, it is less likely that the best fit relationships will share the same exponents of system performance. More so perhaps than interdependency simulation, the interdependency index may be of particular interest to the commercial sector as a simple and easily conceptualised risk management tool. As such, it is important that its derivation and application is straightforward and efficient and this is best achieved by adopting a standardised linear function for all pairs of systems.

5.2.3 Sensitivity tests

The linear function derived in Figure 5.18 is tested for its sensitivity to variations in the availability of back-up power and variations in the fragility of water supply system assets. The existing linear function is the baseline, which is defined as the ‘Medium’ scenario. For each variable four sensitivity tests are conducted representing different levels of reliability: ‘Very low’, ‘Low’, ‘High’ and ‘Very high’ scenarios, with each test applied to the same non-stochastic catalogue used to derive the baseline scenario. For back-up power the four scenarios represent changes in the number of pumping stations that have diesel generators, as shown in Table 5.14, with additions or removals randomly allocated. The same hazard and fragility results from the original analysis are assumed, but at the interdependency simulation stage, new performances

are calculated based on the new allocation of generators. For each back-up power sensitivity test, the fragility parameters of pumping stations and pipes is kept constant at the baseline values.

For the fragility sensitivity tests the four scenarios represent changes in the parameters of the pumping station fragility functions and pipe repair rate functions. For pumping stations, the median value of the baseline fragility function is multiplied by the coefficients in Table 5.14. For pipes, the baseline repair rate function is multiplied by the coefficients in Table 5.14. The same hazard and electric power network results from the original analysis are assumed but the fragility and performance of the water supply system are re-analysed. For each fragility sensitivity test, both the pumping station and pipe fragilities are adjusted while the number of back-up generators remains constant at the baseline value. The best fit linear functions for the sensitivity tests are shown in Figure 5.20 alongside the baseline function, the 'Medium' scenario.

Table 5.14 – Parameters for sensitivity tests for deriving interdependency index from non-stochastic catalogue

Scenario	Back-up generators	Coefficient for pump fragility	Coefficient for pipe repair rate
Very high	92	2	0.25
High	69	1.5	0.5
Medium	46	1	1
Low	23	0.5	1.5
Very low	11	0.25	2

For the sensitivity tests in which the number of back-up generators was adjusted, the plots show a relatively wide spread of results with values for the index ranging from 0.27 to 0.80. The direction of the variation is as expected in that the more reliable scenarios (i.e. more back-up power generators) elicit smaller interdependency responses since there are fewer pumping stations that are vulnerable to power outage.

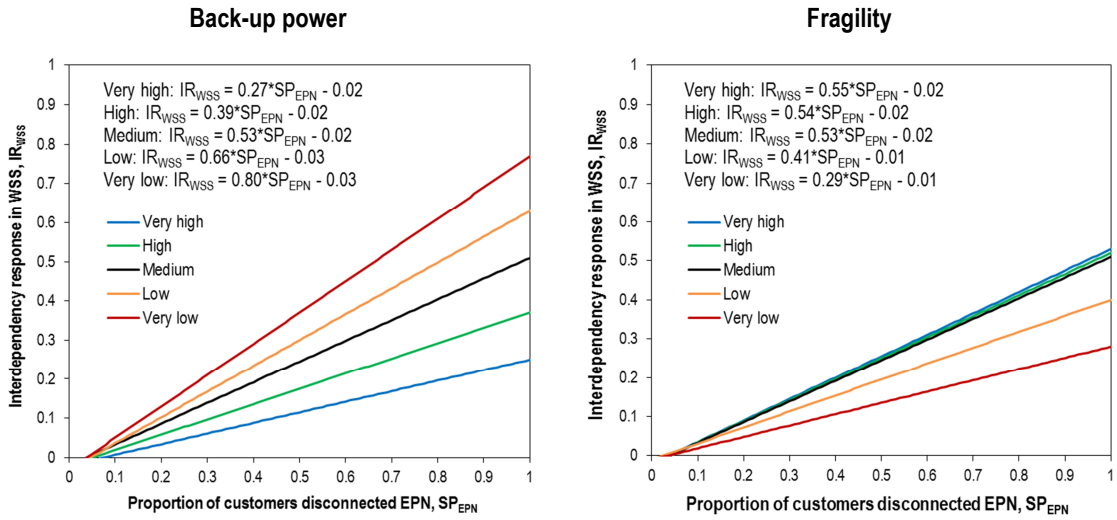

Figure 5.20 – Plots of linear best fit interdependency functions from back-up power (left) and fragility (right) sensitivity tests

For the fragility sensitivity tests however, two interesting phenomena are observed. Firstly, there is almost no difference in the best fit lines for the baseline, 'High' and 'Very high' scenarios. Secondly, the direction of variation is reversed so that it is the less reliable scenarios (i.e. more fragile assets) in which the smaller interdependency responses are observed. The behaviour of the interdependency response metric in relation to asset fragility can be investigated algebraically. The interdependency response between the two systems for the baseline scenario is given by Eq. 5.11. It is defined as the proportion of population whose water supply connection survives direct physical damage but fails due to power outage.

$$I_{WSS|EPN}SP_{EPN} = \frac{SP_{WSS} - SP'_{WSS}}{1 - SP'_{WSS}} = IR_{WSS} \tag{5.11}$$

For an alternative model where the fragility of the water supply assets is adjusted, one would expect the pre-interdependency performance, SP'_{WSS}, to change by some unknown value, Δ_{SP}. Therefore, Eq. 5.11 needs to be modified to account for this. The revised pre-interdependence performance metric, $SP'_{WSS|SENS}$, is simply given by Eq.5.12, where Δ_{SP} is positive if fragility increases.

$$SP'_{WSS|SENS} = SP'_{WSS} + \Delta_{SP} \tag{5.12}$$

In the sensitivity tests where fragility is increased, the post-interdependency performance has to account for the fact that there may be some overlap between the additional population that has become disconnected due to the higher fragility and the population that were already disconnected in the baseline scenario due to interdependency effects. This is shown diagrammatically in Figure 5.21, where the bars represent the proportion of the population disconnected.

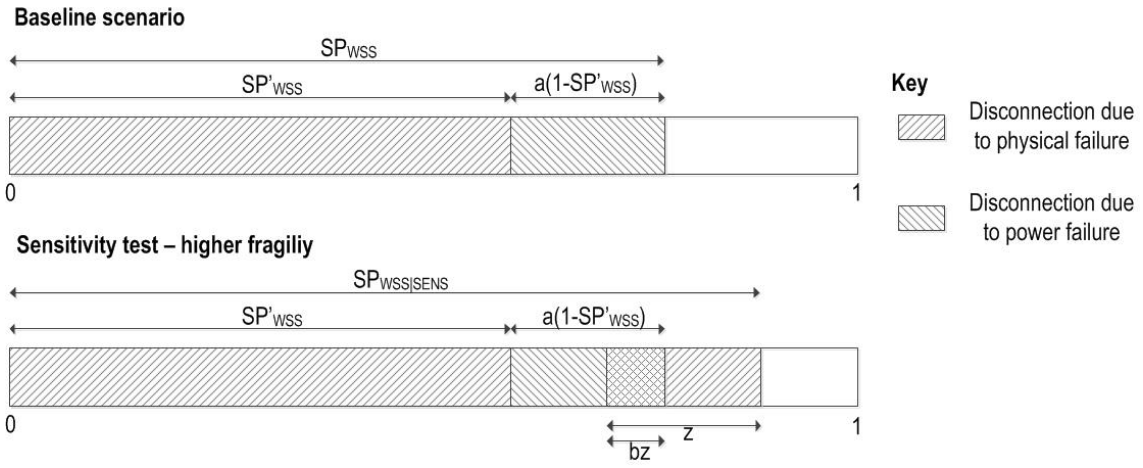

Figure 5.21 – Diagrammatic representation of variables required to calculate interdependency response in baseline scenario and increased fragility sensitivity test

Consequently, to avoid double counting, the revised post-interdependency performance, $SP_{WSS|SENS}$, is given by Eq. 5.13, where ρ_{WSS} is the proportion of the additional disconnected population that overlaps. In the sensitivity tests where fragility is decreased, the post-interdependency performance must account for the fact that some of the population that is no longer disconnected due to the lower fragility, may now become disconnected due to interdependency instead. Eq. 5.13 can still be used to represent this but *z* becomes negative and the definition of ρ_{WSS} changes to represent this new proportion.

$$SP_{WSS|SENS} = SP'_{WSS} + IR_{WSS}\left(1 - SP'_{WSS}\right) + \Delta_{SP} - \rho_{WSS}\Delta_{SP} \tag{5.13}$$

Figure 5.20 shows that for two of the sensitivity tests, there is little or no difference in the interdependency response to the baseline scenario. Each pairwise combination of these three regressions is compared in a Z-test, (as defined in Eq. 5.1), in which the null hypothesis is that the slopes of the two regression are equal. In each case the p-value exceeds 0.99 and so the null hypothesis cannot be rejected, confirming that the regressions are not significantly different. This means that across the three models, for a particular event, the interdependency response is the same regardless of the fragility functions used. This occurs when the condition expressed in Eq. 5.14 is true. Substituting for Eq.'s 5.12 and 5.13 gives Eq. 5.15 and this is rearranged from Eq.'s 5.16 to 5.20 show that if there are two models, in which the only difference is the fragility of pumping stations and pipes, they will produce the same interdependency response if the values of IR_{WSS} and ρ_{WSS} are the same. The variation in fragility, represented by ΔSP, drops out and is not a factor.

$$IR_{WSS} = \frac{SP_{WSS|SENS} - SP'_{WSS|SENS}}{1 - SP'_{WSS|SENS}} \tag{5.14}$$

$$IR_{WSS} = \frac{IR_{WSS}\left(1 - SP'_{WSS}\right) - \rho_{WSS}\Delta_{SP}}{1 - SP'_{WSS} - \Delta_{SP}} \tag{5.15}$$

$$1 = \frac{1 - SP'_{WSS}}{1 - SP'_{WSS} - \Delta_{SP}} - \frac{\rho_{WSS}\Delta_{SP}}{IR_{WSS}\left(1 - SP'_{WSS} - \Delta_{SP}\right)} \tag{5.16}$$

$$\frac{\rho_{WSS}\Delta_{SP}}{IR_{WSS}\left(1 - SP'_{WSS} - \Delta_{SP}\right)} = \frac{1 - SP'_{WSS}}{1 - SP'_{WSS} - \Delta_{SP}} - 1 \tag{5.17}$$

$$\frac{\rho_{WSS}\Delta_{SP}}{IR_{WSS}\left(1 - SP'_{WSS} - \Delta_{SP}\right)} = \frac{\Delta_{SP}}{1 - SP'_{WSS} - \Delta_{SP}} \tag{5.18}$$

$$\rho_{WSS} \Big/ IR_{WSS} = 1 \tag{5.19}$$

$$\rho_{WSS} = IR_{WSS} \quad (5.20)$$

The variables IR_{WSS} and ρ_{WSS} both relate to proportions of customers affected by interdependency. Their values are driven by two aspects: whether or not the pumping station that the customer is connected to has a functioning mains power supply and if not, whether the pumping station has a back-up generator. Neither value appears to be explicitly affected by fragility and so this explains why the baseline scenario and the two more reliable sensitivity tests produce almost identical regressions. Figure 5.20 also shows that for the two less reliable fragility sensitivity tests the interdependency responses are lower. For this to occur, then substituting '=' with '>' in Eq.'s 5.14 to 5.20 gives the condition that ρ_{WSS} must be greater than IR_{WSS}. Given that IR_{WSS} is a function of the baseline pre-interdependency system performance and therefore constant across all sensitivity tests for a specific event, the implication is that the value of ρ_{WSS} must be influenced by fragility. When fragility is low or moderate, $\rho_{WSS} = IR_{WSS}$, but when fragility is high, the value of ρ_{WSS} increases and $\rho_{WSS} > IR_{WSS}$. This can be explained with reference to Figure 5.21. If two models with different fragilities are being compared, but both have low or moderate fragility, then the sum $SP'_{WSS} + \Delta_{SP}$ is low and only affects a small proportion of the population. Assuming that $IR_{WSS,}$ and therefore $IR_{WSS}(1-SP'_{WSS})$ is relatively small, the likelihood of overlap between Δ_{SP} and $IR_{WSS}(1-SP'_{WSS})$ is also relatively small. When fragility is high then the sum $SP'_{WSS} + \Delta_{SP}$ affects a higher proportion of the population and since, for any two models being compared, the terms SP'_{WSS} and $IR_{WSS}(1-SP'_{WSS})$ are constant across both models for a specific event, the likelihood of overlap between Δ_{SP} and is $IR_{WSS}(1-SP'_{WSS})$ is higher. The conclusion is that whilst the availability of back-up power is the primary driver behind the interdependency response, when a system has low physical reliability, it has an indirect influence on interdependency response, since assets that are vulnerable to power loss are already classified as 'failed' before interdependency is considered. Consequently the potential for interdependency to have an effect on the system is reduced.

5.2.4 Extreme event method

A proposed simplified method for deriving the interdependency index, $I_{WSS|EPN}$, is to estimate it from the mean interdependency response in the water supply system for cases where there is complete failure of the electric power network (the extreme event method), as described in section 3.3. Implementation of the extreme event method is conditional on the assumption that the relationship between interdependency response and electric power network performance is linear and passes through, or very close to, the origin – this should occur logically since there should be no interdependency response if there is no loss of performance in the electric power network. The data presented in Figure 5.13 and subsequent analysis in section 5.2.1 confirm that these are reasonable assumptions for this case study, even though a linear function may not be the optimal fit. Whilst for this case study there is little benefit in applying the extreme event method over the slope method, the purpose of this analysis is to determine more generally whether the simpler extreme event method provides a similar result.

The seismic risk assessment in section 5.1 yields just three events in which total loss of performance occurs in the electric power network. The interdependency responses in these three cases are 0.30, 0.45 and 0.78, giving a mean value for the interdependency index of 0.51. This is similar to the value of 0.53 obtained by the slope method but the validity of a value obtained from just three observations is questionable and the similarity of outputs may be coincidental. However, the events in seismic risk assessment catalogue can be artificially manipulated so that a more robust application of the extreme event method is possible. For the manipulated catalogue hazard and water supply system fragility are not re-analysed and the same ground motions, ground deformations and physical damage states are assumed as per the original analysis for each event. However, the fragility of the electric power network is discarded and at the interdependency simulation stage, the electric power network is simply assumed to have suffered complete failure and the effect on pumping station is analysed accordingly. Every event is therefore classified as an extreme event and contributes to the calculation of the interdependency index using this method. Given that the prior analyses show that electric power

network is considerably more reliable to earthquakes than the water supply system, these manipulated events do not represent realistic earthquake scenarios (n.b. they could be realistic representations of other non-seismic hazards, e.g. atmospheric hazards, which are more likely to affect the electric power network than the water supply network due to the presence of overhead cables, or anthropological hazards such as terrorism or cyber-attacks that may be more targeted to the electric power network). As a means of analysing interdependencies, which is not solely a seismic phenomenon, these scenarios are valid. Figure 5.22 shows the frequency distribution of the interdependency indices derived from the 14,428 events in the manipulated catalogue.

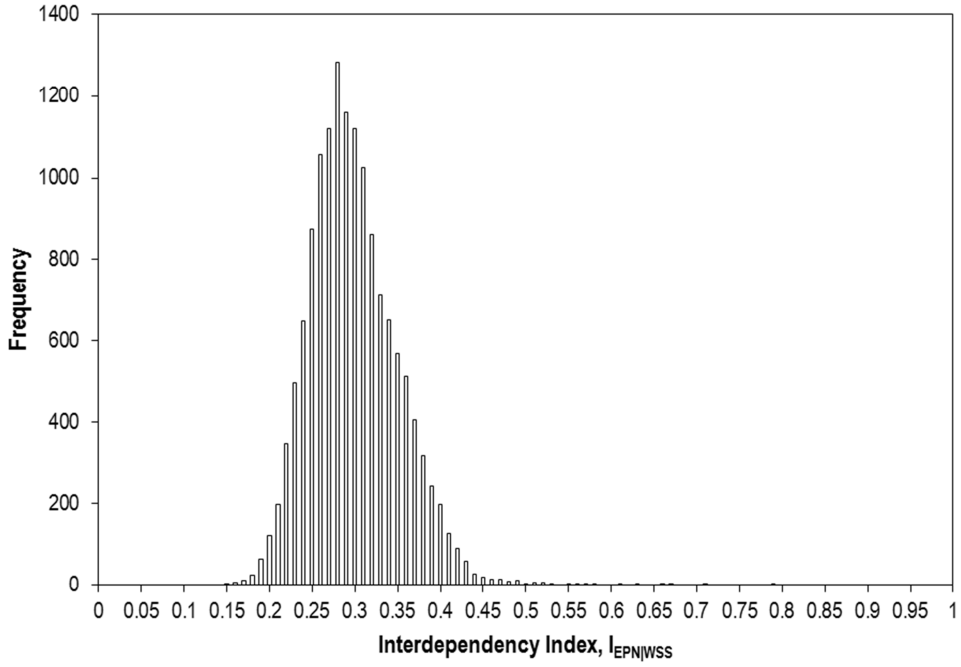

Figure 5.22 – Frequency distribution of interdependency indices derived from manipulated seismic risk assessment catalogue

The results in Figure 5.22 show interdependency indices predominantly in the range between 0.15 and 0.45 and a value of 0.5 is exceeded in only 32 events. The distribution of the results has a coincident mean and median value of 0.29 and a standard deviation of 0.05. Assuming a normal distribution, the 95% confidence interval for the mean is 0.19 to 0.40. The interdependency indices derived using the slope method, 0.48 for the stochastic catalogue and 0.53 for the non-stochastic catalogue, are all significantly higher than this range. In a normal

distribution the probability of an observation more than three standard deviations higher than the mean (i.e. > 0.44) is 0.0015.

Given the upwards kink observed in the interdependency response in Figure 5.18, when the electric power network disconnection is high, it might have been expected that an analysis focused only on high levels of electric power network disconnection would lead to high interdependency responses and therefore a higher index. What is observed to occur instead is that, although barely visible in Figure 5.22, the manipulated catalogue produces values of interdependency response ranging from 0.15 to 0.78, whilst the original stochastic catalogue produces responses ranging from 0.30 to 0.78 when there is total disconnection in the electric power network, albeit with only three observations. In the non-stochastic catalogue there are no events with total disconnection of the electric power network but there are 17 events in which disconnection is 0.99. Amongst these events, the interdependency responses range from 0.36 to 0.76 and if the extreme event method is applied to this data, it would yield an interdependency index of 0.61.The median response in the manipulated catalogue is 0.29, meaning that over half the observed responses in the manipulated catalogue are lower than the range of responses observed for the 17 events in the non-stochastic catalogue in which disconnection is 0.99. From an interdependency perspective, the manipulation of the stochastic catalogue has introduced a significant number of low impact events that skew the average response downwards.

This skewing is a result of the difference in the nature of the hazard in the catalogues. The original stochastic and non-stochastic catalogues contain only realistic earthquake scenarios, whereas the manipulated catalogue contains a much wider range of performance scenarios including many that would have an extremely low probability of occurring due to an earthquake but that may occur due to other hazards. The reason why realistic earthquakes produce higher interdependency responses can be demonstrated by investigating the relationship between the electric power network performance and the pre-interdependency performance of the water supply system. Figure 5.23 shows the two performances plotted against each other per event for the original stochastic and non-stochastic catalogues.

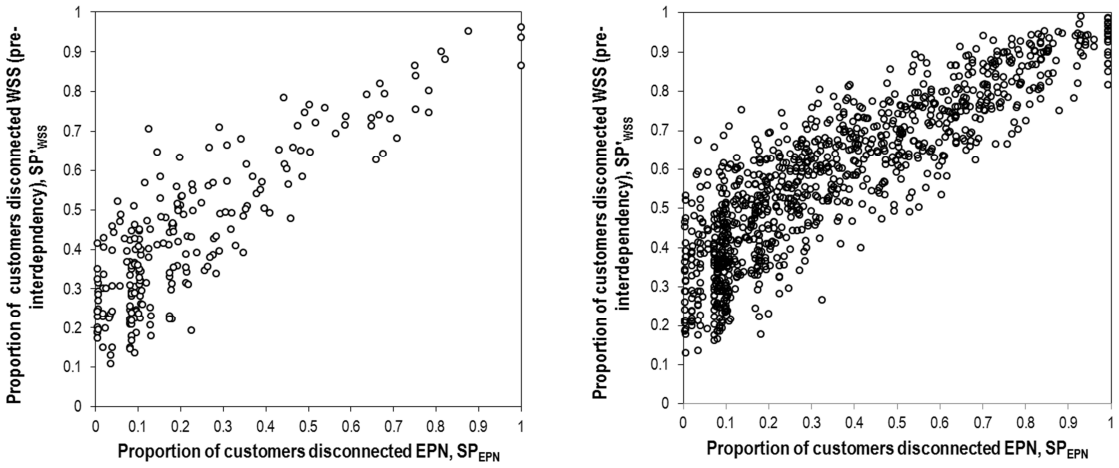

Figure 5.23 – Scatter plot of performance metrics observation pairs for the electric power network and water supply system for the original stochastic (left) and non-stochastic (right) catalogues

In both cases there is a strong positive correlation between the two performances, 0.85 for the stochastic catalogue and 0.87 for the non-stochastic catalogue. The plots show that when disconnection in the electric power network is high, disconnection in the water supply system is also high. In the non-stochastic catalogue, for events where disconnection in the electric power network exceeds 0.9 (38 events), the minimum level of disconnection in the water supply system is 0.78. In the manipulated catalogue, while disconnection in the electric power network is fixed at 1.0, the disconnection in the water supply system is distributed as shown in Figure 5.23, where less than 1% of events result in disconnection greater than 0.5.

Furthermore, Figure 5.24 shows how the interdependency response varies with pre-interdependency performance of the water supply system for the original stochastic and non-stochastic catalogues. Although the relationship is not as strong as with the electric power network, the two plots do suggest a moderate trend for interdependency response increasing as pre-interdependency disconnection increases. Therefore, in realistic earthquake scenarios, both electric power network disconnection and pre-interdependency water supply system disconnection are expected to be high at the same time due to observed correlation. Furthermore, since both these conditions are related to high values of interdependency response, it follows that in realistic earthquake scenarios, interdependency response will be high. In the manipulated catalogue where the correlation between performances has been artificially fixed to

zero, the inclusion of events in which the pre-interdependency disconnection in the water supply system is low causes the downwards shift in the average interdependency response.

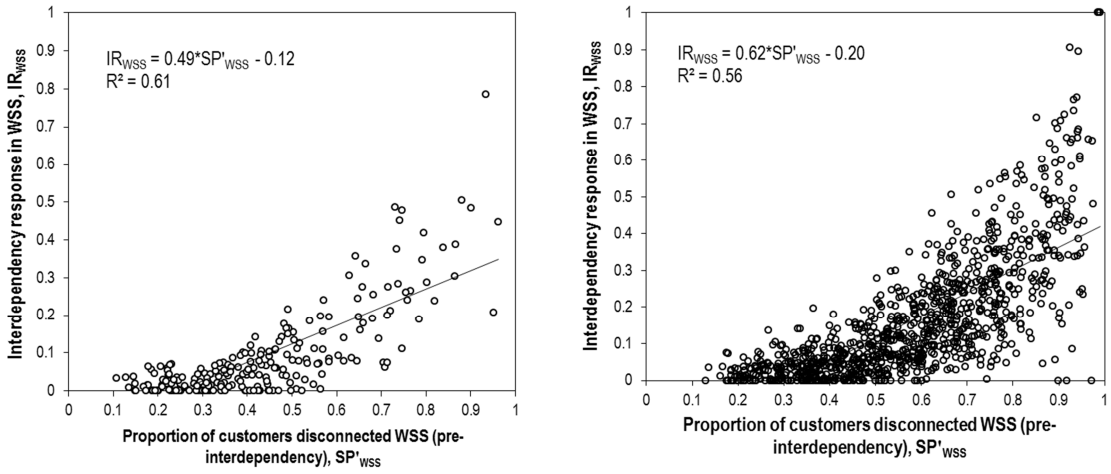

Figure 5.24 – Plot of observed interdependency response against pre-interdependency water supply system performance for 228 events in stochastic catalogue where electric power network suffered loss of performance (left) and non-stochastic catalogue (right)

The analysis indicates that the scale of the dependency of one system on another varies according to the type of hazard and the type of the damage and performance scenarios that can realistically be expected from that hazard. The use of a manipulated or artificial catalogue, in which the performance of at least one of the systems is assumed rather than explicitly modelled, ignores any correlation between performances that might occur due to a specific type of hazard. It is representative of a wide range of possible performance combinations irrespective of hazard. An index derived in this way could therefore be described as hazard-independent. However, for hazards such as earthquakes where there is correlation between performances, only a narrower range of performance combinations is likely. Therefore for an analysis involving only this hazard it is more accurate to derive an index based only on realistic scenarios, (although this may only be possible when resources are available for explicit fragility and performance modelling of both systems).

The proposed extreme event method relies on there being events in which there is total disconnection in the supporting system. In this case study, neither of the realistic catalogues produced a large number of applicable events. To approximate the interdependency index,

relaxing the extreme event criterion to disconnection levels above 0.9, yields 38 events in the non-stochastic catalogue with an average interdependency response of 0.59. This is higher than calculated by the slope method (0.53) but is based on a considerably smaller sample so the validity of the extreme event method as an approximation of the slope method is inconclusive. The lack of observations acts as a limiting factor to the application of the extreme event method to realistic catalogues. In this case study where the supporting system is highly reliable, there are too few events with high levels of disconnection to make the extreme event method viable. However, there may be more potential for application of the method to case studies where the supporting system is less physically reliable.

6. Conclusions

6.1 Research summary

The literature review in Chapter 2 demonstrates that although the seismic risk assessment of critical infrastructure systems is not a new topic for academic research, there are shortcomings in previous studies, particularly with respect to the lack of calibration or validation of proposed models and the treatment of interdependencies. Although there is now around 15 years of literature on the issue interdependencies, many studies still fail to consider them and there is no consensus on their application in seismic risk assessment. There is also no consensus on their measurement, with many studies subjectively quantifying the strength of coupling between systems. More specifically for the case of insurance loss estimation, there are no existing methods or frameworks that are completely adequate for commercial application and there are no guidelines for the development of new seismic risk assessment models. The findings of the literature review support the three objectives of the research set out in the introduction, which are: to produce a modelling framework that outlines methodologies for the development and application of seismic risk assessment models for the specific requirements of the insurance sector; to demonstrate the application of the modelling framework with a case study validated on observed data; and to propose and test a method for the quantification of interdependencies between infrastructure systems.

A general method for conducting the seismic risk assessment of critical infrastructure systems that is applicable to the insurance sector is proposed in Chapter 3. The method brings together individual concepts from existing literature such as Hazus and the SYNER-G project and is framed in terms of the three loss estimation modules that are already widely in use in catastrophe models in the insurance sector: hazard, vulnerability and loss. The modules are sub-divided into a simple eight-step procedure: the classification of infrastructure system assets (exposure); the creation of a stochastic event catalogue; the prediction of seismic hazard intensities; the evaluation of asset failure probabilities; the deterministic assignment of

functional states to assets; the measurement of system performance accounting for interdependencies; the modelling of asset restoration times; and the estimation of direct and indirect financial losses. The proposed seismic risk assessment method is only conceptual but is followed by a set of more specific model development guidelines that will enable insurers to select model components that are tailored to their study area of interest rather than relying on a one-size-fits-all model that has itself been tailored to a specific geographical region (e.g. Hazus for the United States, SYNER-G for Europe). The framework allows for the development of new model components and functions based on empirical data if these are available or the adoption of existing model components from the literature, provided their use can be validated and justified. The guidelines acknowledge the limitations that insurers have in terms of data availability, resources and expertise for model development and so in the case of some of the most complex or poorly understood model components – e.g. spatial correlation, liquefaction triggering and interdependency modelling – the guidelines lists specific methods that are suitable within these limitations.

The model development framework is applied to the case study of the electric power network and the water supply system in Christchurch. As well as being a demonstration of the framework, the case study produces significant research value through the development of new model components and functions to illuminate problems that have not been adequately addressed previously. In relation to seismic hazard, the case study leads to the development of a new spatial correlation model and a new simplified liquefaction triggering model. A range of existing methods for developing spatial correlation models are tested using data from strong motion stations recorded during the Canterbury earthquake sequence and the model developed using the method of Jayaram and Baker (2009) is found to be the most efficient. It is the first spatial correlation model that has been developed for Christchurch (or for anywhere in New Zealand), and can be applied to any future study that requires the prediction of ground motions in the city and surrounding area.

Three existing methods for simplified modelling of liquefaction triggering are tested including the Liquefaction Potential Index (LPI), Hazus and the Zhu et al. (2015) model. Comparisons of model predictions with observations from the Darfield and Christchurch earthquakes using contingency table analysis show that the LPI method performs best. Two new modifications to the LPI method are also proposed, which make use of V_{S30} data, which is more widely available than the V_S data needed for the original method. Whilst the model based on direct substitution of V_S for V_{S30} does not yield good predictions, the model that uses of V_{S30} to simulate realistic V_S profiles is the second best performing model. Given the issues relating to data availability, the proposed modified LPI method (with simulated V_S profiles) will be more useful to the insurance industry as a tool for assessment of liquefaction triggering. All three versions of the model proposed by Zhu et al. (2015) also perform well, but with the caveat that the level of performance by applying a very low threshold probability (0.1) for liquefaction occurrence, suggesting that the model may be biased towards negative prediction, although it must be acknowledged that the Zhu et al. models have not been developed with the intention of making site-specific predictions. It is found that the Hazus model performs very poorly, with significant bias towards negative prediction, i.e. underestimating the extent of liquefaction. This is an important result, since although the Hazus method is only intended for the United States, it is used widely by researchers for studies located elsewhere and has effectively become the default method for simplified liquefaction triggering analysis regardless of location. The analysis in this thesis shows not only that the assumption of the Hazus model potentially unreliable outside of the United States, but also that there are other alternative simplified methods that could be applied instead. In the proposed seismic risk assessment model development guidelines, it is advised that all simplified models are tested in order to find the one most suitable for the study area.

For damage assessment, new fragility and repair rate functions are developed for buried cables, substations, pipes, wells and pumping stations. The repair rate functions developed for buried cables are the first to ever be produced that distinguish buried cables from overhead lines. They

are developed for specific typologies that exist in the Christchurch distribution network (PILCA, XLPE, PILCA HDPE), but the paucity of alternatives for modelling damage to electrical conduits means that they may be of value for any analysis of seismic damage to buried cables. The repair rate analysis finds that insulation material is the primary factor affecting the fragility of cables, which is relatively unaffected by conduction material or age. The analysis also supports the hypothesis of Kwasinski et al. (2014), who postulated that liquefaction is the predominant hazard affecting cable performance and that ground shaking should have minimal impact, since the materials used to construct most electrical cables should be able to absorb the levels of ground strain generated by earthquakes. For substations, two existing methods (Hazus and SYNER-G) are tested against observations from the Darfield and Christchurch earthquakes and both are found to significantly overestimate damage. Since there are insufficient data to produced completely new fragility functions, Bayesian analysis is applied using the GEM procedure to update the Hazus fragility functions with the observed data. The resulting fragility function, which only predicts probability of failure, is applicable beyond Christchurch to substations in any system that has undergone significant seismic strengthening.

Although there are many existing repair rate functions in the literature, this thesis has produced new repair rate functions specifically for application in Christchurch. These differ from other repair rate functions developed for Christchurch by O'Rourke et al. (2014) since they are developed using a higher resolution analysis of observed IMs. The functions are developed for specific typologies that exist in Christchurch and more generally for brittle and ductile pipes. For wells and pumping stations, Hazus is the only existing source for fragility functions. Comparison of model predictions with observations from the Darfield and Christchurch earthquakes shows that the Hazus function significantly overestimates damage. Therefore, Bayesian analysis is applied using the GEM procedure to update the Hazus fragility functions with the observed data, to create a new function for estimating probability of failure. Since there is a lack of existing fragility functions for modelling damage to these assets, the new functions

are useful as they provide a possible alternative to the original Hazus functions for any study on seismic performance of water supply systems.

A selection of candidate models are applied to a simplified version of the electric power network and water supply system in Christchurch and validated by comparing predicted system performance metrics with corresponding observed values. Two methods for predicting liquefaction triggering (LPI and modified LPI with simulated V_S profiles) and two methods for applying interdependencies (by mapped supply zones and by geographical proximity) are tested in the candidate models. The original LPI method leads to significant under-prediction of customer disconnection in the electric power network, whereas the mean performance prediction when using the modified LPI method has an error (relative to observed performance) that is comparable to previous critical infrastructure risk assessments. With the constraint of applying the modified LPI methods, the smallest error in the prediction of water supply system performance is achieved when interdependencies are designated by supply zones. This error is also comparable to previous studies. It is also noted however, that the modelling combination of the original LPI method and interdependencies designated by geographical proximity also produces low error predictions for the water supply system. No wider conclusions can therefore be drawn regarding the effectiveness of the liquefaction triggering or interdependency modelling methods, but for this specific case study, a model that uses the modified LPI method and designates interdependencies by supply zones is the only one that is valid. The validated model is tested for geospatial precision in the electric power network by comparing outage predictions to observations at the meshblock level. For the water supply system, geospatial precision is assessed by comparison of predicted to observed proportions of pipe repairs in each pressure zone. Both tests find that the model is sufficiently precise in its geospatial predictions of damage.

The final validated model is applied for future risk projection in Christchurch. In hazard terms, the model generates 14,428 events in 10,000 years. Despite the devastating effects of events in the Canterbury earthquake sequence, the model indicates that liquefaction is a rare occurrence,

which is predicted to occur in less than 3% of events. As expected, the model predicts that the areas the River Avon are likely to be most frequently affected by liquefaction, while the Port Hills area to the south of city is least vulnerable to liquefaction. However, it is the predicted engineering risk that is of most value. The forecasts from the engineering analysis can be used by the system operators, Orion and Christchurch City Council, to assist with disaster management and emergency planning. In particular, the results help to identify assets most at risk of failure, and neighbourhoods most at risk of disconnection.

The analysis shows that the electric power network is extremely reliable in earthquakes, with damage only predicted in 228 events and in most cases the level of customer disconnection is low (<25%). This reflects the observation that buried cables are predominantly affected by liquefaction rather than ground shaking and is a vindication of the strategy implemented by Orion to seismically strengthening their assets following the Edgecumbe earthquake. The locations of electric power network assets that are most frequently predicted to fail correspond to the areas that are also most vulnerable to liquefaction, around the city centre and the River Avon, with the Oxford-Tuam and Montreal substations at particularly high risk. However, sensitivity testing of potential mitigation strategies identify the Dallington, Pages-Kearneys and Rawhiti substations as those that are most beneficial to system performance if strengthened. Geospatial predictions of outage by meshblock show that the most areas at most risk are the city centre, the suburbs to the north and east and a small part of the Port Hills area to the south. However, whilst the electric power network is reliable in most events, the system risk curve shows that significant disruption, including potentially complete outage, can be expected for strong low frequency events. This is an important finding since there is a high dependency on electric power from other critical infrastructure systems, in particular water supply, waste water collection, storm water collection, telecommunications, railways and air transport (Lamb, 1997). Therefore, these low frequency events can be considered to be high impact events in terms of infrastructure reliability.

The water supply system is significantly less reliable than the electric power network with loss of performance predicted to occur in all events although the level of customer disconnection is mostly between 10%-30%. Damage to pipes is predicted in every event including weaker high frequency events. This indicates that seismic upgrading of pipes should be a priority action. Pumping stations are more reliable with failures predicted to occur in only 500 events, although in the most damaging earthquakes, a higher proportion of pumping stations are predicted to fail than pipes. Therefore, whilst the fragility of pumping stations, and in particular wells, which are representative of primary pumping station fragility, become a critical factor in strong long return period earthquakes, the main source of vulnerability in the water supply system across all events is the fragility of the pipes. Sensitivity testing of potential mitigation strategies shows that strengthening of even the most vulnerable pumping stations leads to negligible improvement in system reliability. Geospatial predictions of disconnection by meshblock show that the areas at most risk are Riccarton, Edgeware, Linwood, Halswell, Lyttelton and Governor's Bay. System risk curves produced for the two interdependency methods confirm that modelling interdependency by proximity is conservative compared to modelling interdependency by known connections, although due to the reliability of the electric power network the difference is only observable in low frequency events.

Two methods for quantifying interdependency using results from a seismic risk assessment model are proposed in this research, based on modifications to the Leontief input-output method. For both methods two new metrics are defined. The interdependency response is the proportional loss of functionality in a system due specifically to interdependency effects and varies between events. The interdependency index is the expected interdependency response in a dependent system when there is total failure in the supporting system, and is assumed to be constant for a pair of systems. The interdependency index takes a value between 0 and 1 and is an objective measure of interdependency that can be used for risk management, as proposed in the research objectives.

In the first method (the ‘linear slope method’), a linear function is fit through a plot of predicted interdependency responses in the water supply system against predicted electric power network system performances and the gradient of the best fit line is approximately equal to the interdependency index. The analysis is conducted for the 228 events from the future risk projection in which outages are predicted for the electric power network and then repeated for a non-stochastic event set designed to increase the number of observations at poor performance levels. Both analyses produce plots with a high correlation and for which a linear function is a good fit, and hence validates the proposed method in concept. The stochastic events yield an interdependency index of 0.48 and the non-stochastic events yield an interdependency index 0.53, both of which indicate a moderate level of interdependency between the two systems. However, the slopes are found to be significantly different and it is concluded that the result from the non-stochastic event set method is more reliable for this analysis, since it incorporates a better distribution of events (in terms of system performance) and in any case, the proposed interdependency index is deterministic and so does not require a probabilistic event set. If this method is applied by others, it is recommended that a non-stochastic catalogue is used to avoid underestimating the interdependency index. Non-linear functions are also tested and do produce better fits through the data. However, there are significant complexities associated with application of indices derived in this way and comparisons of indices across different pairs of systems are invalid if they have been derived using different functional forms. Therefore this method is not suitable for commercial application.

Sensitivity tests are applied to the ‘linear slope method’ to assess how the interdependency index is affected by the availability of back-up power and by the fragility of assets in the water supply system. The results show that the interdependency index changes relatively uniformly with the availability of back-up power but in the opposite direction, i.e. increasing the number of generators decreases the index. Asset fragility only affects the index when fragility is high. In this scenario, the increase in the number of assets that are physically damaged means that there are fewer assets remaining that are susceptible to interdependent failure. Whilst increasing

fragility leads to a smaller interdependency index, decreasing fragility has negligible impact. Therefore, although seismic upgrading of assets improves the physical reliability of a system, it potentially has no impact on the ability of the system to resist the effects of interdependencies.

The second proposed interdependency quantification method (the 'extreme event method') is a simplification of the 'linear slope method' that takes advantage of the assumption that the best fit line in the 'linear slope method' passes close to the origin and therefore, its slope is equal to the expected interdependency response in the dependent system when there is total outage in the supporting system. The analysis shows that this can be estimated from the results of a seismic risk assessment of the two systems, as long as there are sufficient data at that performance level. However, if performance data are available for both systems, there is little value in this method since the 'linear slope method' can be applied with only marginally greater computational effort. The potential value of the 'extreme event method' derives from being able to estimate the interdependency index when performance data are only available for the dependent system. This is done by artificially fixing the performance in the supporting system to total outage for every scenario event. However, the analysis shows that this method is only valid for hazards where the performance of the two systems is uncorrelated and so the artificial performance combinations are realistic. For earthquakes, there is likely to be high correlation between system performances and so the method cannot be used. Therefore, the proposed 'extreme event method' may have some limited value for quantifying interdependency, but not in the specific case of seismic risk.

6.2 Discussion of limitations

Although this research has produced a number of important and useful findings of scientific interest, it is necessary to highlight some of its limitations. New liquefaction triggering models have been developed for Christchurch based on USGS V_{S30} data as a proxy for V_S profiles. However, this is only example and no conclusions can be drawn on the method's wider applicability Furthermore since the method is based on V_S approximations, it is recommended to only use it for regional scale analysis and not for site-specific analysis. It is stated in the model

development framework that an analyst should not develop an empirical model for predicting the scale of permanent ground deformations. The lack of existing models however, leads to a reliance on the Hazus models, which have been developed specifically for the United States and may not be valid elsewhere. This is a significant issue for infrastructure assets such as buried cables and pipes, which are particularly susceptible to liquefaction. Despite this, no attempt has been made in this research to develop a new prediction model.

The fragility analysis and damage assignment methods proposed in the seismic risk assessment method focus on operational failure rather than damage states. This is due to the focus of this research on system performance (to inform indirect loss estimations), which only requires knowledge of whether an asset is operational. For comprehensive loss estimation however, modelling of damage states and subsequent repair costs should be carried out. More detailed system performance modelling could also be conducted by considering serviceability in addition to connectivity. Connectivity analysis may overstate the true performance of a system as it ignores scenarios where a connection exists but the level of service is sub-standard. However, consideration of serviceability adds significant complexity to the analysis and is not deemed to be suitable for application in the insurance sector.

There are limitations in relation to specific components of the case study model. In the fragility analysis, only empirical methods are considered for the development of new fragility functions. This is because the model development guidelines have been produced primarily for use in the insurance sector, for whom empirical methods are the most practical, and the case study application is intended to reflect that. In practice however, it is possible that analytical methods would produce more accurate fragility functions for the assets that are studied. Also in relation to fragility functions, no uncertainty in measurements of ground shaking or ground deformation are considered, although it should be noted that this is exceptionally rare amongst existing fragility functions and so does not depart from current practice. This may however be particularly significant for the fragility and repair rate functions developed in terms of permanent ground deformations, given the uncertainty associated with the LiDAR

measurements used here. Restoration time modelling is important for predicting business interruption losses, which are time dependent. However, it has not been considered in the case study due to limitations associated with existing functions and a lack of observed data with which to develop new functions. The procedure for simulating interdependencies by nodal analysis, becomes more complex when there are more than two systems, especially if there are feedback relationships. Although the procedure for dealing with this is briefly alluded to in the literature review, it has not been possible to demonstrate this practically due to a lack of data from additional infrastructure systems.

As part of the model development testing, the case study draws some conclusions regarding the suitability of different methods, specifically in relation to liquefaction triggering, fragility functions and interdependency simulation. It must be stressed that these conclusions are valid only for the Christchurch case study. More generally, The final model has not been validated on independent data, and this could be considered to be a limitation. However, it is important to consider that independent validations are non-existent in the literature on post-disaster critical infrastructure performance and even examples such as this case study, where models are developed and tested on the same dataset are exceptionally rare. Therefore, although the lack of independent validation is in theory a limitation, when viewed within the context of previous studies, the exercise presented in this thesis exceeds most current practice.

For interdependency quantification, the proposed methods have been based on the Leontief input-output method for interdependency simulation. This has been chosen for its relative simplicity and the clarity it gives to the definition of interdependency indices. As such it is the most suitable approach for insurers or risk managers who require a simple and objective method for describing interdependency. It has been assumed that whilst there other conceptual approaches to interdependency modelling (e.g. system dynamics, agent-based modelling) that could be used as a basis for quantification, these are too complex for the intended end use and so this research has not attempted to use these methods. Furthermore, the proposed interdependency quantification methods are applied only to measurements of system

performance in terms of customer disconnection, which are numerically constrained. No other metrics are tested and so no conclusions can be drawn regarding the wider applicability of the methods. Also the case study to which the methods are applied considers only two systems with a one-way dependency. Due to lack of data, it does not consider the case where a system is dependent on more than one supporting system, or the case where there is a two-way dependency that causes feedback. The former is straightforward but requires additional modelling steps, since 'marginal' system performances need to be modelled to isolate the individual effect of each supporting system so that distinct plots of interdependency response versus supporting system performance can be produced. The latter is more complex since the feedback introduces multiple cycles of observations of interdependency response versus supporting system performance and so further investigation is needed to consider how data from different cycles are treated.

6.3 Future research opportunities

The limitations of this thesis also present opportunities for potential future research. The model development guidelines have only been applied to a single case study and it would valuable to demonstrate its utility by applying it to other case studies where sufficient data are available. It would be of particular value if they could be applied to a case study where there are data for empirical model development and independent data for model validation. Given that infrastructure performance data is generally poorly recorded after earthquakes, the likelihood of two high quality independent datasets for the same location is low. However, as the earthquake engineering, insurance and disaster management communities become increasingly conscious of the seismic risk to infrastructure, there is an opportunity for practitioners to start taking a more proactive approach to data collection in disaster scenarios, rather than relying on infrastructure operators whose focus is usually elsewhere in the aftermath of natural catastrophes. Other valuable extensions to the case study presented here would be the inclusion of more than two systems, the inclusion of systems with multiple dependencies, and the inclusion of two-way dependencies with feedback. Further comparative studies would also be useful to test the wider

validity of the conclusions drawn in this research concerning the relative performance of interdependency simulation by geographical proximity and interdependency simulation by known connections.

Liquefaction is extremely important to critical infrastructure risk and the prediction of permanent ground deformations is a major shortcoming that needs to be addressed. In terms of liquefaction triggering, this research proposes a modified version of the LPI method based on simulated profiles. While it is shown to predict liquefaction relatively well in Christchurch, its wider applicability needs to be tested with a new case study. A comparison between the effectiveness of the empirical fragility functions proposed in this research and alternative fragility functions developed by analytical methods would be useful to assist Orion and Christchurch City Council with the risk management of their assets. Restoration time modelling is important for estimating business interruption losses and for emergency management purposes but there are few existing sources for this. Additional studies into restoration times and how they interact with interdependencies have the potential to provide significant commercial value in the catastrophe modelling sector.

This research proposed two methods for interdependency quantification. The ‘linear slope method’ depends on being able to approximate the plot of interdependency responses versus supporting system performances as a linear function that passes close to the origin. This was demonstrated for the case study presented here but should be tested with new case studies. There is also a need to test the method with more complicated interdependency scenarios, such as dependencies on multiple systems or two-way interdependencies. The ‘extreme event method’ does not work with seismic risk due to the correlation of performances between systems. It may be applicable for scenarios where performances are uncorrelated but this hypothesis needs to be tested.

6.4 Concluding remarks

Despite the limitations that have been identified and the gaps in research that have yet to be explored, the investigation presented in this thesis makes a significant contribution to both academic and commercial understanding of seismic risk to critical infrastructure. The spatial correlation model for Christchurch will enhance future studies of seismic hazard in the city and the new fragility functions will enable more precise modelling of risk to infrastructure assets in Christchurch, but also more widely they present an alternative to the limited number of existing functions and especially the often automatic adoption of Hazus functions. In particular the repair rate function for buried cables will be extremely valuable since it is the first of its kind globally. The model development framework will be useful for then insurance sector for the development of bespoke models and the simplified method that has been developed for liquefaction triggering will enhance the capability of insurers to model this hazard, which is currently deficient. Finally, the analysis of interdependency response in the water supply system has provided new theoretical insight into the relationship between dependent and supporting infrastructure systems, and the proposed interdependency index, which is objective and straightforward to derive, is a useful tool for insurers or other analysts to quantify those relationships and understand their relative importance for risk management.

References

Adachi, T. and Ellingwood, B.R. (2008). Serviceability of earthquake-damaged water systems: effects of electrical power availability and power back-up systems on system vulnerability. *Reliability Engineering and System Safety*, 93(1), 78-88.

Adachi, T. and Ellingwood, B.R. (2009). Serviceability assessment of a municipal water system under spatially correlated seismic intensities. *Computer-Aided Civil and Infrastructure Engineering*, 24(4), 237-248.

Adachi, T. and Ellingwood, B.R. (2010). Comparative assessment of civil infrastructure network performance under probabilistic and scenario earthquakes. *Journal of Infrastructure Systems*, 16(1), 1-10.

Agresti, A. (2002). *Categorical Data Analysis*. 2nd ed. Hoboken, NJ: Wiley Interscience.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.

Akkar, S. and Bommer, J.J. (2007). Empirical prediction equations for peak ground velocity derived from strong-motion records from Europe and the Middle East. *Bulletin of the Seismological Society of America*, 97(2), 511-530.

Akkar, S. and Bommer, J.J. (2010). Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean region and the Middle East. *Seismological Research Letters*, 81(2), 195-206.

AIR. (2012). *About Catastrophe Models*. Boston: AIR Worldwide.

AIR. (2014). *AIR Earthquake Model for the United States*. Boston: AIR Worldwide.

AIR. (2016). *AIR Estimates Insured Losses from M7.0 Japan Earthquake at Between JPY 80 billion (USD 1.7 billion) and JPY 320 billion (USD 2.9 billion)*. Retrieved from http://www.air-worldwide.com/Press-Releases/AIR-Estimates-Insured-Losses-from-M7-0-Japan-Earthquake-at-Between-JPY-180-Billion-(USD-1-7-Billion)-and-JPY-320-Billion-(USD-2-9-Billion)/

ALA. (2001). *Seismic Fragility Formulations for Water Systems: Part 1 – Guideline*. American Lifelines Alliance.

ALA. (2005). *Guideline for Assessing the Performance of Electric Power Systems in Natural Hazard and Human Threat Events*. American Lifelines Alliance (ALA).

Anagnos, T. and Ostrom, D.K. (2000). Electrical substation equipment damage database for updating fragility estimates. In *12th World Conference on Earthquake Engineering, 30 January – 4 February 2000, Auckland, New Zealand*.

Andrus, R.D. and Stokoe, K.H. (2000). Liquefaction resistance of soils from shear-wave velocity. *Journal of Geotechnical and Geoenvironmental Engineering*, 126(11), 1015-1025.

Ang, A.H.-S., Pires, J.A. and Villaverde, R. (1996). A model for the seismic reliability assessment of electric power transmission systems. *Reliability Engineering and System Safety*, 51, 7-22.

ATC. (1985). *ATC-13 Earthquake Damage Evaluation Data for California*. Redwood City, California: Applied Technology Council (ATC).

ATC. (1991). *ATC-25 Seismic Vulnerability and Impact of Disruption of Lifelines in the Conterminous United States*. Redwood City, California: Applied Technology Council (ATC).

Azevedo, J., Guerreiro, L., Bento, R., Lopes, M. and Proença, J. (2007). Seismic vulnerability of lifelines in the greater Lisbon area. *Bulletin of Earthquake Engineering*, 8(1), 157-180.

Baker, J. (2015). Efficient analytical fragility function fitting using dynamic structural analysis. *Earthquake Spectra*, 31(1), 579-599.

Bakun, W.H. (1984). Seismic moments, local magnitude and coda-duration magnitudes for earthquakes in central California. *Bulletin of the Seismological Society of America*, 74, 439-485.

Bazzurro, P. and Cornell, C.A. (1999). Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America*, 89(2), 501-520.

Beaven, J., Samsonov, S., Motagh, M., Wallace, L., Ellis, S. and Palmer, N. (2010). The Darfield (Canterbury) earthquake: geodetic observations and preliminary source model. *Bulletin of the New Zealand Society for Earthquake Engineering*, 43(4), 228-235.

Berahman, F. and Behnamfar, F. (2007). Seismic fragility curves for un-anchored on-grade steel storage tanks: Bayesian approach. *Journal of Earthquake Engineering*, 11(2), 166-192.

Bensi, M., Straub, D., Friis-Hansen, P. and Der Kiureghian, A. (2009). In H. Furuta, D. Frangopol, M. Shinozuka (Eds.), *Proceedings of the 10th International Conference on Structural Safety and Reliability, ICOSSAR, 13-17 September 2009, Osaka, Japan*. CRC Press.

Bensi, M., Der Kiureghian, A., and Straub, D. (2013). Efficient Bayesian network modeling of systems. *Reliability Engineering and System Safety*, 112, 200-213.

Bigger, J.E., Willingham, M.G. and Krimgold, F. (2009). Consequences of critical infrastructure interdependencies: lessons from the 2004 hurricane season in Florida. *International Journal of Critical Infrastructures*, 5(3), 199-219.

Billings, I. and Charman, N. (2011). Christchurch City Council lifelines – performance of concrete potable water reservoirs in the February 2011 Christchurch earthquake. *Bulletin of the New Zealand Society for Earthquake Engineering*, 44(4), 418-424.

Bird, J.F. and Bommer, J.J. (2004). Earthquake loss due to ground failure. *Engineering Geology*, 75(2), 147-179.

Bird, J.F., Bommer, J.J., Crowley, H. and Pinho, R. (2006). Modelling liquefaction-induced building damage in earthquake loss estimation. *Soil Dynamics and Earthquake Engineering*, 26(1), 15-30.

Bolt, B.A. (2006). *Earthquakes*. 5th Ed. New York: W.H. Freeman and Company.

Bommer, J.J. (2002). Deterministic vs probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy. *Journal of Earthquake Engineering*, 6(S1), 43-73.

Bonneau, A.L. and O'Rourke, T.D. (2009). *Water Supply Performance During Earthquakes and Extreme Events, Technical Report MCEER-09-0003*, Buffalo, NY: Multidisciplinary Center for Earthquake Engineering Research (MCEER).

Boore, D.M. (2004). Estimating $V_S(30)$ (or NEHRP site classes) from shallow velocity models (depths < 30 m). *Bulletin of the Seismological Society of America*, 94(2), 591-597.

Boore, D.M., Gibbs, J.F., Joyner, W.B., Tinsley, J.C. and Ponti, D.J. (2003). Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard bridge collapse, west Los Angeles, California. *Bulletin of the Seismological Society of America*, 93(6), 2737-2751.

Botts, H., Du, W., Foust, B., Ihinger, P. and Jeffery, T. (2012). *Modeling Earthquake Risk*. Irvine, CA: CoreLogic.

Brackley, H.L. (2012). *Review of Liquefaction Hazard Information in Eastern Canterbury, Including Christchurch City and Parts of Selwyn, Waimakariri and Hurunui Districts*. Christchurch, New Zealand: Environment Canterbury.

Bradley, B.A. (2013). A New Zealand-specific pseudospectral acceleration ground-motion prediction equation for active shallow crustal earthquakes based on foreign models. *Bulletin of the Seismological Society of America*, 103(3), 1801-1822.

Bradley, B.A. and Cubrinovski, M. (2011). Near-source strong ground motions observed in the 22 February 2011 Christchurch earthquake. *Bulletin of the New Zealand Society for Earthquakwe Engineering*, 44(4), 181-194.

Bradley, B.A. and Hughes, M.W. (2012). *Conditional Peak Ground Accelerations in the Canterbury Earthquakes for Conventional Liquefaction Assessment*. Christchurch, New Zealand: Department of building and Housing.

Brillinger, D.R. (1993). Earthquake risk and insurance. *Environmetrics*, *4*(1), 1–21.

Bruneau, M., Chang, S.E., Eguchi, R.T., Lee, G.C., O'Rourke, T.D., Reinhorn, A.M., Shinozuka, M., Tierney, K., Wallace, W.A. and von Winterfeldt, D. (2003). A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthquake Spectra*, *19*(4), 733–752.

Buxton, R., Uma, S.R. and King, A.B. (2010). Modelling interdependencies of critical infrastructures. In *2010 New Zealand Society for Earthquake Engineering Conference, 26-28 March 2010, Wellington, New Zealand*.

Cagnan, Z., Davidson, R.A. and Guikema, S.D. (2006). Post-earthquake restoration planning for Los Angeles electric power. *Earthquake Spectra*, 22(3), 589-608.

Campbell, K.W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in Eastern North America. *Bulletin of the Seismological Society of America*, 93(3), 1012-1033.

Cavalieri, F., Franchin, P., Buriticá Cortés, J.A.M. and Tesfamariam, S. (2014a). Models for seismic vulnerability analysis of power networks: comparative assessment. *Computer-Aided Civil and Infrastructure Engineering*, 29, 590-607.

Cavalieri, F., Franchin, P. and Pinto, P.E. (2014b). Fragility functions of electric power stations. In K. Pitilakis, H. Crowley, A. Kaynia (Eds.), *SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk*. Dordrecht, Netherlands: Springer.

Cavalieri, F., Franchin, P. and Pinto, P.E. (2014c). Application to Selected Transportation and Electric Networks in Italy. In K. Pitilakis, P. Franchin, B. Khazai and H. Wenzel (Eds.), *SYNER-G: Systemic Seismic Vulnerability and Risk Assessment of Complex Urban, Utility, Lifeline Systems and Critical Facilities*. Dordrecht, Netherlands: Springer.

CCC. (2005). *Water and Sanitary Services Assessment*. Christchurch, New Zealand: Christchurch City Council (CCC).

CERA. (2013). *CERA/CERA_ZoneMapAndTechClass (MapServer)*. Retrieved from http://maps.cera.govt.nz/arcgis/rest/services/CERA/CERA_ZoneMapAndTechClass/MapServer

Chang, L. and Song, J. (2007). Matrix-based system reliability analysis of urban infrastructure networks: a case study of MLGW natural gas network. In *5th China-Japan-US Trilateral Symposium on Lifeline Earthquake Engineering*.

Chang, S.E. (1996). Economic impact of lifeline disruption in the January 17, 1995 Hanshin-Awaji earthquake. In *11th World Conference on Earthquake Engineering, 23-28 June 1996, Acaculpo, Mexico*.

Chang, S.E. and Nojima, N. (2001). Measuring post-disaster transportation system performance: the 1995 Kobe earthquake in comparative perspective. *Transportation Research Part A*, 35, 475-494.

Chang, S.E., Svekla, W.D. and Shinozuka, M. (2002). Linking infrastructure and urban economy: simulation of water-disruption impacts in earthquakes. *Environment and Planning B: Planning and Design*, 29(2), 281-301.

Chang, S.E., McDaniels, T.L., Mikawoz, J. and Peterson, K. (2007). Infrastructure failure interdependencies in extreme events: power outage consequences in the 1998 Ice Storm. *Natural Hazards*, 41(2), 337-358.

Charvet, I., Ioannou, I., Rossetto, T., Suppasri, A. and Imamura, F. (2014). Empirical fragility assessment of buildings affected by the 2011 Great East Japan tsunami using improved statistical models. *Natural Hazards*, 73, 951-973.

Chiou, B. and Youngs, R.R. (2008). An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, 24(1), 173-215.

Chiou, B., Youngs, R.R., Abrahamson, N. and Addo, K. (2010). Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in California and its implications on regionalization of ground-motion prediction models. *Earthquake Spectra*, 26(4), 907-926.

Choe, D.-O., Gardoni, P., Rosowsky, D. and Haukaas, T. (2009). Seismic fragility estimates for reinforced concrete bridges subject to corrosion. *Structural Safety*, 31(4), 275-283.

Cimellaro, G.P. (2013). Resilience-based design (RBD) modelling of civil infrastructure to assess seismic hazards. In S. Tesfamariam and K. Goda (Eds.), *Handbook of Seismic Risk Analysis and Management of Civil Infrastructure Systems*. Cambridge, United Kingdom: Woodhead Publishing.

Clark, K.M. (1986). A formal approach to catastrophe risk assessment and management. *Proceedings of the Casualty Actuarial Society*, *73*(2), 69–92.

Clark, K.M. (2002). The use of computer modeling in estimating and managing future catastrophe losses. *The Geneva Papers on Risk and Insurance*, *27*(2), 181–195.

Coburn, A. and Spence, R. (2002). *Earthquake Protection* (2nd ed.). London: Wiley.

Cornell, C.A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5), 1583-1606.

D'Ayala, D., Meslem, A., Vamvatsikos, D., Porter, K., Rossetto, T., Crowley, H. and Silva, V. (2013). *Guidelines for Analytical Vulnerability Assessment*. Pavia, Italy: Global Earthquake Model Foundation.

Daniell, J.E., Vervaeck, A., Khazai, B. and Wenzel, F. (2012). Worldwide CATDAT damaging earthquakes database in conjunction with earthquake-report.com - presenting past and present socio-economic earthquake data. In *15th World Conference on Earthquake Engineering, 24-28 September 2012, Lisbon, Portugal*.

Daniell, J.E., Khazai, B., Wenzel, F. and Vervaeck, A. (2011). The CATDAT damaging earthquakes database. *Natural Hazards and Earth System Science*, *11*(8), 2235–2251.

Daniell, J.E., Wenzel, F. and Khazai, B. (2010). The cost of historic earthquakes today – economic analysis since 1900 through the use of CATDAT. In *Australian Earthquake Engineering Society 2010 Conference, 26-28 November 2010, Perth, Australia*.

Davis, M.W. (1987). Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Mathematical Geology*, 19(2), 91-98.

Delavaud, E., Cotton, F., Akkar, S., Scherbaum, F., Danciu, L., Beauval, C., Drouet, S., Douglas, J., Basili, R., Abdullah Sandikkaya, M., Segou, M., Faccioli, E. and Theodoulidis, N. (2012). Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. *Journal of Seismology*, 16(3), 451-473.

De Smith, M.J., Goodchild, M.F. and Longley, P.A. (2009). *Geospatial Analysis: a Comprehensive Guide to Principles, Techniques and Software Tools*. Winchelsea, UK: The Winchelsea Press.

DHS. (2013). *NIPP 2013: Partnering for Critical Infrastructure Security and Resilience*. Washington, DC: Department for Homeland Security (DHS).

Dizhur, D., Ismail, N., Knox, C., Lumantarna, R. and Ingham, J.M. (2010). Performance of unreinforced and retrofitted masonry buildings during the 2010 Darfield earthquake. *Bulletin of the New Zealand Society for Earthquake Engineering*, 43(4), 321-339.

Dobson, A.J., Kuulasmaa, K., Eberle, E. and Scherer, J. (1991). Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine*, 10, 457-462.

Douglas, J. (2015). *Ground motion prediction equations 1964-2015*.Retrieved from http://www.gmpe.org.uk/

Drayton, M.J. and Verdon, C.L. (2013). Consequences of the Canterbury earthquake sequence for insurance loss modelling. In *2013 New Zealand Society for Earthquake Engineering Conference, 26-28 April 2013, Wellington, New Zealand*.

Dudenhoeffer, D.D., Permann, M.R., Woolsey, S., Timpany, R., Miller, C., McDermott, A. and Manic, M. (2007). Interdependency modeling and emergency response. In *Summer Computer Simulation Conference, 16-19 July 2007, San Diego, CA*.

Dueñas-Osorio, L., Craig, J.I. and Goodno, B.J. (2004). Probabilistic response of interdependent infrastructure networks. In *2nd Annual Meeting of the Asian-Pacific Network of Centers for Earthquake Engineering Research*.

Dueñas-Osorio, L., Craig, J.I. and Goodno, B.J. (2007). Seismic response of critical interdependent networks. *Earthquake Engineering and Structural Dynamics*, 36, 285-306.

Dueñas-Osorio, L. and Kwasinski, A. (2012). Quantification of lifeline system interdependencies after the 27 February 2010 M_W 8.8 offshore Maule, Chile, earthquake. *Earthquake Spectra*, 28(S1), 581-603.

Earthquake Commission. (2016). *The New Zealand Geotechnical Database*. Retrieved from https://www.nzgd.org.nz/Registration/Login.aspx?ReturnUrl=%2f

ECan. (2014a). *Consent Search*. Retrieved from http://ecan.govt.nz/services/online-services/pages/consent-search.aspx

ECan. (2014b). *Canterbury Liquefaction Susceptibility*. Retrieved from http://canterburymaps.co.nz/Map/?webmap=e47bcfe622f146328b222d8ac12114f9#.UzoEhbUyaPs

Eidinger, J. and Tang, A.K. (2012). *Christchurch, New Zealand Earthquake Sequence of MW 7.1 September 04, 2010, MW 6.3 February 22, 2011, MW 6.0 June 13, 2011: Lifeline Performance, Technical Council on Lifeline Earthquake Engineering Monograph No. 40*. Reston, VA: American Society of Civil Engineers.

Esposito, S. (2011). *Systemic seismic risk analysis of gas distribution networks* (Unpublished doctoral thesis). University of Naples Federico II, Naples, Italy.

Esposito, S. and Iervolino, I. (2011). PGA and PGV spatial correlation models based on European multievent datasets. *Bulletin of the Sesimological Society of America*, 101(5), 2532-2541.

Esposito, S., Iervolino, I., d'Onofrio, A., Santo, A., Cavalieri, F. and Franchin, P. (2015). Simulation-based seismic risk assessment of gas distribution networks. *Computer-Aided Civil and Infrastructure Engineering*, 30(7), 508-523.

Eusgeld, I., Nan, C. and Dietz, S. (2011). "System-of-systems" approach for interdependent critical infrastructures. *Reliability Engineering and System Safety*, 96(6), 679-686.

FEMA. (1999). *Hazus Earthquake Loss Estimation Methodology, Technical Manual*. Washington D.C.: Federal Emergency Management Agency (FEMA).

FEMA. (2003). *Multi-Hazard Loss Estimation Methodology, Earthquake Model, Hazus®-MH, Technical Manual*. Washington D.C.: Federal Emergency Management Agency (FEMA).

FEMA. (2015). *Multi-Hazard Loss Estimation Methodology, Earthquake Model, Hazus®-MH 2.1, Technical Manual*. Retrieved from https://www.fema.gov/media-library/assets/documents/24609

Foulser-Piggott, R. and Stafford, P. (2010). Incorporation of the spatial correlation of Arias intensity within earthquake loss estimation. In *5th International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, 24-29 May 2010, San Diego, United States*.

Franchin, P. (2014). A Computational Framework for Systemic Risk Analysis of Civil Infrastructural Systems. In K. Pitilakis, P. Franchin, B. Khazai and H. Wenzel (Eds.), *SYNER-G: Systemic Seismic Vulnerability and Risk Assessment of Complex Urban, Utility, Lifeline Systems and Critical Facilities*. Dordrecht, Netherlands: Springer.

Friedman, D.G. (1984). Natural hazard risk assessment for an insurance program. *The Geneva Papers on Risk and Insurance*, *9*(30), 57–128.

Fujisaki, E., Takhirov, S., Xie, Q. and Mosalam, K.M. (2014). Seismic vulnerability of power supply: lessons learned from recent earthquakes and future horizons of research. In A. Cunha, E. Caetano, P. Ribeiro and G. Muller (Eds.), *Proceedings of the 9th International Conference on Structural Dynamics, EURODYN 2014, 30 June – 2 July 2014, Porto, Portugal*. Porto, Portugal: Faculty of Engineering of the University of Porto.

Galli, E. (2010). *Agent based modeling and simulation for critical and interdependent systems* (Unpublished doctoral thesis). Universita di Roma “Tor Vergata”.

Gardoni, P. and Rosowsky, D. (2009). Seismic fragility increment functions for deteriorating reinforced concrete bridges. *Structure and Infrastructure Engineering*, 7(11), 869-879.

Gehl. P., Desramaut, N., Réveillere, A. and Modaressi, H. (2014). Fragility functions of gas and oil networks. In K. Pitilakis, H. Crowley, A. Kaynia (Eds.), *SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk*. Dordrecht, Netherlands: Springer.

Giovinazzi, S., Wilson, T., Davis, C., Bristow, D., Gallagher, M., Schofield, A., Villemure, M., Eidinger, J. and Tang, A. (2011). Lifelines performance and managementn following the 22 February 2011 Christchurch earthquake, New Zealand: highlights of resilience. *Bulletin of the New Zealand Society for Earthquake Engineering*, 44(4), 402-417.

Goda, K. and Atkinson, G.M. (2009). Probabilistic characterization of spatially corelated response spectra for earthquakes in Japan. *Bulletin of the Seismological Society of America*, 99(5), 3003-3020.

Goda, K. and Hong, H.P. (2008). Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America*, 98(1), 354-365.

Goda, K., Atkinson, G.M., Hunter, A.J., Crow, H. and Motazedian, D. (2011). Probabilistic liquefaction hazard analysis for four Canadian cities. *Bulletin of the Seismological Society of America*, 101(1), 190-201.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. New York: Oxford University Press.

Gordon, P., Richardson, H.W. and Davis, B. (1998). Transport-related impacts of the Northridge earthquake. *Journal of Transportation and Statistics*, 1(2), 21-36.

GNS Science. (2014). *GeoNet Quake Search*. Retrieved from http://quakesearch.geonet.org.nz/

Green, R.A., Cubrinovski, M., Cox, B., Wood, C., Wotherspoon, L., Bradley, B.A. and Maurer, B. (2014). Select liquefaction case histories from the 2010-2011 Canterbury earthquake sequence. *Earthquake Spectra*, 30(1), 131-153.

Grossi, P., Kunreuther, H. and Windeler, D. (2005). An Introduction to Catastrophe Models and Insurance. In P. Grossi and H. Kunreuther (Eds.), *Catastrophe Modeling*. Boston: Springer Verlag.

Grossi, P. and Kunreuther, H. (2005). *Catastrophe Modeling*. Boston: Springer Verlag.

Grossi, P. and Windeler, D. (2005). Sources, Nature and Impact of Uncertainties on Catastrophe Modeling. In P. Grossi and H. Kunreuther (Eds.), *Catastrophe Modeling*. Boston: Springer Verlag.

Grossi, P. and TeHennepe, C. (2008). A Guide to Catastrophe Modelling. In G. Dobie (Ed.), *The Review - Worldwide Reinsurance: A Guide to Catastrophe Modelling*. London: Informa.

Guha-Sapir, D., Below, R. and Hoyois, P. (2015). EM-DAT: International Disaster Database. Retrieved February 22, 2016, from http://www.emdat.be

Guikema, S. and Gardoni, P. (2009). Reliability estimation for networks of reinforced concrete bridges. *Journal of Infrastructure Systems*, 15(2), 61-69.

Gutenberg, B. and Richter, C.F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34, 185-188.

Haimes, Y.Y. and Jiang, P. (2001). Leontief-based model of risk in complex interconnected infrastructures. *Journal of Infrastructure Systems*, 7(1), 1-12.

Han, S.Y. and DeLaurentis, D. (2013). Development interdepedency modeling for system-of-systems (SoS) using Bayesian networks: SoS management strategy planning. *Procedia Computer Science*, 16, 698-707.

Hernandez-Fajardo, I. and Dueñas-Osorio, L. (2011). Sequential propagation of seismic fragility across interdependent lifeline systems. *Earthquake Spectra*, 27(1), 23-43.

Hijmans, R., Williams, E. and Vennes, C. (2015). *Package 'geosphere'*. Retrieved from https://cran.r-project.org/web/packages/geosphere/vignettes/geosphere.pdf

Hines, W.W., Montgomery, D.C., Goldsman. D.M., and Borror, C.M. (2003). *Probability and Statistics in Engineering*. 4th ed. Hoboken, NJ: John Wiley and Sons Inc.

Hosmer, D.W. and Lemeshow, S. (1980). A goodness-of-fit test for the multiple logistic regression model. *Communications in Statistics*, A10, 1043-1069.

Hughes, M.W., Quigley, M.C., van Ballegooy, S., Deam, B.L., Bradley, B.A., Hart, D.E. and Measures, R. (2015). The sinking city: earthquakes increase flood hazard in Christchurch, New Zealand. *Geology*, 25(3), 4-10.

Hwang, H.H.M. and Huo, J.-R. (1998). Seismic fragility analysis of electric substation equipment and structures. *Probabilistic Engineering Mechanics*, 13(2), 107-116.

Hwang, H.H.M., Lin, H. and Shinozuka, M. (1998). Seismic performance assessment of water delivery systems. *Journal of Infrastructure Systems*, 4, 118-125.

Iervolino, I., Fabbrocino, G. and Manfredi, G. (2004). Fragility of standard industrial structures by a response surface based method. *Journal of Earthquake Engineering*, 8(6), 927-945.

Isoyama, R., Ishida, E., Yune, K. and Shirozu, T. (2000). Seismic damage estimation procedure for water supply pipelines. In *12th World Conference on Earthquake Engineering, 30 January – 4 February 2000, Auckland, New Zealand*.

Iwasaki, T., Arakawa, T. and Tokida, K.-I. (1984). Simplified procedures for assessing soil liquefaction during earthquakes. *Soil Dynamics and Earthquake Engineering*, 3(1), 49-58.

Jaiswal, K., Wald, D. and D'Ayala, D. (2011). Developing empirical collapse fragility functions for global building types. *Earthquake Spectra*, 27(3), 775-795.

Javanbarg, M.B. and Takada, S. (2009). Seismic reliability assessment of water supply systems. In H. Furuta, D. Frangopol, M. Shinozuka (Eds.), *Proceedings of the 10th International Conference on Structural Safety and Reliability, ICOSSAR, 13-17 September 2009, Osaka, Japan*. CRC Press.

Jayaram, N. and Baker, J.W. (2009). Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics*, 38, 1687-1708.

Jayaram, N. and Baker, J.W. (2010). Efficient sampling and data reduction techniques for probabilistic seismic lifelines risk assessment. *Earthquake Engineering and Structural Dynamics*, 39, 1109-1131.

Juang, C.H., Yang, S.H. and Yuan, H. (2005). Model uncertainty of shear wave velocity-based method for liquefaction potential evaluation. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(10), 1274-1282.

Kakderi, K. and Argyroudis, S. (2014). Fragility functions of water and waste-water systems. In K. Pitilakis, H. Crowley, A. Kaynia (Eds.), *SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk*. Dordrecht, Netherlands: Springer.

Kaklamanos, J., Boore, D.M., Thompson, E.M. and Campbell, K.W. (2010). *Implementation of the Next Generation Attenuation Ground Motions (NGA) Ground Motion Prediction Equations in Fortran and R, US Geological Survey Open-File Report 2010-1296*, Reston, VA: US Geological Survey.

Kaklamanos, J., Baise, L.G. and Boore, D.M. (2011). Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. *Earthquake Spectra*, 27(4), 1219-1235.

Kam, W.Y., Pampanin, S. and Elwood, K. (2011). Seismic performance of reinforced concrete buildings in the 22 February Christchurch (Lyttelton) earthquake. *Bulletin of the New Zealand Society for Earthquake Engineering*, 44(4), 239-278.

Kameda, H. (2000). Engineering management of lifeline systems under earthquake risk. In *12th World Conference on Earthquake Engineering, 30 January – 4 February 2000, Auckland, New Zealand*.

Kanamori, H. and Anderson, D.L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, 65, 1073-1095.

Kang, W.-H., Song, J. and Gardoni, P. (2008). Matrix-based system reliability method and application to bridge networks. *Reliability Engineering and System Safety*, 93(11), 1584-1593.

Kang, W.-H., Lee, Y.-J., Song, J. and Gencturk, B. (2012). Further development of matrix-based system reliability method and applications to structural systems. *Structure and Infrastructure Engineering*, 8(5), 441-457.

Kawakami, H. (2000). Earthquake performance of highway system in Tokyo. In *12th World Conference on Earthquake Engineering, 30 January – 4 February 2000, Auckland, New Zealand.*

Kim, Y., Song, J., Spencer, B.F. and Elnashai, A.S. (2009). Seismic risk assessment of complex interacting infrastructures using matrix-based system reliability method. In In H. Furuta, D. Frangopol, M. Shinozuka (Eds.), *Proceedings of the 10th International Conference on Structural Safety and Reliability, ICOSSAR, 13-17 September 2009, Osaka, Japan*. CRC Press

King, A., Middleton, D., Brown, C., Johnston, D. and Johal, S. (2014). Insurance: its role in recovery from the 2010 – 2011 Canterbury earthquake sequence. *Earthquake Spectra, 30*(1), 475–491.

King, S.A., Kiremidjian, A.S., Pachakis, D. and Sarabandi, P. (2004). Application of empirical fragilityfunctions from recent earthquakes. In *13th World Conference on Earthquake Engineering, 1-6 August 2004, Vancouver, Canada*.

Kircher, C.A., Reitherman, R.K., Whitman, R.V. and Arnold, C. (1997). Estimation of earthquake losses to buildings. *Earthquake Spectra*, 13(4), 703-720.

Kiremidjian, A.S., Moore, J., Fan, Y.Y., Yazlali, O., Basoz, N. and Williams, M. (2007). Seismic risk assessment of transportation network systems. *Journal of Earthquake Engineering*, 11(3), 371-382.

Kozlowski, R.T. and Mathewson, S.B. (1997). A primer on catastrophe modeling. *Journal of Insurance Regulation*, *15*(3), 322–341.

Kruschke, J.K. (2011). *Doing Bayesian Data Analysis: a Tutorial with R and BUGS*. Cambridge, MA: Academic Press.

Kunreuther, H. (2008). Reducing losses from catastrophic risks through long-term insurance and mitigation. *Social Research*, *75*(3).

Kwasinski, A., Eidinger, J., Tang, A.K. and Tudo-Bornarel, C. (2014). Performance of electric power systems in the 2010-2011 Christchurch, New Zealand, earthquake sequence. *Earthquake Spectra*, 30(1), 205-230.

Lakervi, E. and Holmes, E. (1995). *Electricity Distribution Network Design*. Stevenage, UK: The Institution of Engineering and Technology.

Lamb, J. (1997). *Risks and Realities: a Multi-Disciplinary Approach to the Vulnerability of Lifelines to Natural Hazards*. Chrsitchurch, New Zealand: The New Zealand Centre for Advanced Engineering.

Lee, R. and Kiremidjian, A. (2007). Uncertainty and correlation for loss assessment of spatially distributed systems. *Earthquake Spectra*, 23(4), 753- 770.

Leelardcharoen, K. (2011). *Interdependent Response of Telecommunication and Electric Power Systems to Seismic Hazard* (Unpublished doctoral dissertation). Georgia Institute of Technology.

Leontief, W.W. (1986). *Input-Output Economics*. 2nd ed. New York: Oxford University Press.

Leprince, S., Barbot, S., Ayoub, F. and Avouac, J.-P. (2007). Automatic and precise orthorectification, coregistration and subpixel correlation of satellite images, application to ground deformation measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6), 1529-1558.

Liao, S.S, Veneziano, D. and Whitman, R.V. (1988). Regression models for evaluating liquefaction probability. *Journal of Geotechnical Engineering*, 114(4).

LiDAR-UK. (2016). *What is LiDAR?*. Retrieved from http://www.lidar-uk.com/what-is-lidar/

Lloyds. (2015). *Solvency II and Risk Assurance: Guidance Notes*. London.

LMA. (2013). *Catastrophe Modelling: Guidance for Non-Catastrophe Modellers*. London: Lloyd's Market Association (LMA).

Lunn, D., Spiegelhalter, D., Thomas, A. and Best, N. (2009). The BUGS project: evolution, critique and future directions (with discussion). *Statistics in Medicine*, 26, 3049-3082.

McDonald, J.H. (2014). *Handbook of Biological Statistics*. 3rd ed. Baltimore: Sparky House Publishing.

McGinty, P. (2001). Preparation of the New Zealand earthquake catalogue for a probabilistic seismic hazard analysis. *Bulletin of the New Zealand Society for Earthquake Engineering*, 34(1), 60-67.

Mahdyiar, M. and Porter, B. (2005). The Risk Assessment Process: The Role of Catastrophe Modeling in Dealing with Natural Hazards. In P. Grossi and H. Kunreuther (Eds.), *Catastrophe Modeling*. Boston: Springer Verlag.

Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozome. *Biochimica et Biophysica Acta (BBA) – Protein Structure*, 405(2), 442-451.

Maurer, B.W., Green, R.A., Cubrinovski, M. and Bradley, B.A. (2014). Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. *Journal of*

Geotechnical and Geoenvironmental Engineering, 140(7), doi: 10.1061/(ASCE)GT.1943-5606.0001117.

Min, H.-S.J., Beyeler, W., Brown, T., Son, Y.J. and Jones, A.T. (2007). Toward modeling and simulation of critical national infrastructure interdependencies. *IIE Transactions*, 39(1), 57-71.

Ministry of Health. (2012). *Annual Report on Drinking Water Quality 2010-2011*. Wellington, New Zealand: Ministry of Health.

Moschonas, I.F., Kappos, A.J., Panetsos, P., Papadopoulos, V., Makarios, T. and Thanopoulos, P. (2009). Seismic fragility curves for Greek bridges: methodology and case studies. *Bulletin of Earthquake Engineering*, 7(2), 439-468.

Nateghi, R., Guikema, S.D. and Quiring, S.M. (2011). Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes. *Risk Analysis*, 31(12), 1897-1906.

Neighbors, C.J., Cochran, E.S., Caras, Y. and Noriega, G.R. (2013). Sensitivity analysis of FEMA HAZUS earthquake model: case study from King County, Washington. *Natural Hazards Review*, 14, 134-146.

NIWA. (2015). *RiskScape Technical Documentation Wiki*. Retrieved from https://wiki-riskscape.niwa.co.nz/index.php/RiskScape_Technical_Documentation_Wiki

O'Rourke, M.J. and So, P. (2000). Seismic fragility curves for on-grade steel tanks. *Earthquake Spectra*, 16(4), 801-815.

O'Rourke, T.D., Stewart, H.E. and Jeon, S.-S. (2001). Geotechnical aspects of lifeline engineering. *Geotechnical Engineering*, 149(1), 13-26.

O'Rourke, T.D., Jeon, S.-S., Toprak, S., Cubrinovski, M. and Jung, J.K. (2012). Underground lifeline system performance during the Canterbury earthquake sequence. In *15th World Conference on Earthquake Engineering, 24-28 September 2012, Lisbon, Portugal*.

O'Rourke, T.D., Jeon, S.-S., Toprak, S., Cubrinovski, M., Hughes, M.W., van Ballegooy, S. and Bouziou, D. (2014). Earthquake response of underground pipeline networks in Christchurch, NZ. *Earthquake Spectra*, 30(1), 183-204.

Oliva, G., Panzieri, S. and Setola, R. (2010). Agent-based input-output interdependency model. *International Journal of Critical Infrastructure Protection*, 3(2), 76-82.

Orion. (2009). *Asset Management Plan: a 10-year Management Plan for Orion's Electricity Network from 1 April 2009 to 31 March 2019*. Christchurch, New Zealand: Orion Ltd.

Orion. (2014). *Asset Management Plan: a 10-year Management Plan for Orion's Electricity Network from 1 April 2013 to 31 March 2023*. Christchurch, New Zealand: Orion Ltd.

Osiadacz, A. (1987). *Simulation and Analysis of Gas Networks*. Houston: Gulf Publishing Company.

Ottonelli, D., Cattari, S. and Lagomarsino, S. (2015). Urban risk assessment: fragility functions for masonry buildings. In *Proceedings of the 2015 International Conference on Civil Engineering, CIVILENG2015, 16-20 July 2015, Zakynthos, Greece*.

Padgett, J.E. and DesRoches, R. (2008). Methodology for the development of analytical fragility curves for retrofitted bridges. *Earthquake Engineering and Structural Dynamics*, 37, 1157-1174.

Park, J., Nojima, N. and Reed, D.A. (2006). Nisqually earthquake electric utility analysis. *Earthquake Spectra*, 22(2), 491-509.

Paul, P., Pennell, M.L. and Lemeshow, S. (2013). Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Statistics in Medicine*, 32, 67-80.

Pederson, P., Dudenhoeffer, D.D., Hartley, S. and Permann, M.R. (2006). *Critical Infrastructure Interdependency Modeling: a Survey of US and International Research*. Idaho Falls: Idaho National Laboratory.

Pineda-Porras, O. and Najafi, M. (2010). Seismic damage estimation for buried pipelines: challenges after three decades of proogress. *Journal of Pipeline Systems Engineering*, 1, 19-24.

Pineda-Porras, O. and Ordaz, M. (2010). Seismic fragility formulations for segmented buried pipeline systems including the impact of differential ground subsidence. *Journal of Pipeline Systems Engineering*, 1, 141-146.

Pitilakis, K., Alexoudi, M., Argyroudis, S., Monge, O. and Maryin, C. (2006). Earthquake risk assessment of lifelines. *Bulletin of Earthquake Engineering*, 4(4), 365-390.

Pitilakis, K., Franchin, P., Khazai, B. and Wenzel, H. (2014a). *SYNER-G: Systemic Seismic Vulnerability and Risk Assessment of Complex Urban, Utility, Lifeline Systems and Critical Facilities*. Dordrecht, Netherlands: Springer.

Pitilakis, K., Crowley, H. and Kaynia, A. (2014b). *SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk*. Dordrecht, Netherlands: Springer.

Pitilakis, K. and Kakderi, K. (2011). Seismic risk assessment and management of lifelines, utilities and infrastructures. *In 5th International Conference on Earthquake Geotechnical Engineering, 10-13 January 2011, Santiago, Chile.*

Platt, R. (1991). Lifelines: an emergency management priority for the United States in the 1990s. *Disasters*, 15(2), 172-176.

Poljansek, K., Bono, F., Gutiérrez, E. (2012). Seismic risk assessment of interdependent critical infrastructure systems: the case of European gas and electricty networks. *Earthquake Engineering and Structural Dynamics*, 41, 61-79.

Porter, K.A. (2003). An overview of PEER's performance-based earthquake engineering methodology. In A. der Kiureghian, S. Madanat and J. Pestana (Eds.), *Proceedings of the 9th International Conference on Applications of Statistics and Probability in Civil Engineering*. Amsterdam: IOS Press.

Porter, K.A. and Sherrill, R. (2007). Utility performance panels in the ShakeOut scenario. *Earthquake Spectra*, 27(2), 443-458.

Powers, D.M.W. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.

Quigley, M.C., Bastin, S. and Bradley, B.A. (2013). Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence. *Geology*, 41(4), 419-422.

Railsback, S.F. and Grimm, V. (2011). *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton: Princeton University Press.

Rasulo, A., Goretti, A. and Nuti, C. (2004). Performance of lifelines during 2002 Molise, Italy, earthquake. *Earthquake Spectra*, 20(S1), S301-S314.

Rasulo, A., Vanzi, I. and Nuti, C. (2008). Seismic reliability analysis of utility networks. In *15th World Conference on Earthquake Engineering, 12-17 October 2008, Beijing, China*.

Rauch, A.F. and Martin, J.R. (2000). EPOLLS model for predicting average displacements on lateral spreads. *Journal of Geotechnical and Geoenvironmental Engineering*, 126, 360-371.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969-1982. *Journal of Geophysical Research*, 90(B7), 5479-5495.

Reed, D.A., Kapur, K.C. and Christie, R.D. (2009). Methodology for assessing the resilience of networked infrastructure. *IEEE Systems Journal*, 3(2), 174-180.

Reiter, L. (1990). *Earthquake Hazard Analysis: Issues and Insights*. New York: Columbia University Press.

Reitherman, R. (1985). A review of earthquake damage estimation methods. *Earthquake Spectra*, *1*(4), 805-847.

Rinaldi, S.M. (2004). Modeling and simulating critical infrastructures and their interdependencies. In *Proceedings of the 37th Hawaii International Conference on System Sciences, 5-8 January 2004, Kauai, HI*.

Rinaldi, S.M., Peerenbom, J.P. and Kelly, T.K. (2001). Identifying, understanding and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, 21(6), 11-25.

RMS. (2016). *RMS Estimates Industry Loss from Kumamoto, Japan Earthquakes Will Range Between 800 Million and 1.2 Billion USD*. Retrieved from https://www.rms.com/newsroom/press-releases/press-detail/2016-05-11/rms-estimates-industry-loss-from-kumamoto-japan-earthquakes-will-range-between-800-million-and-12-billion-usd

Romero, N., O'Rourke, T.D., Nozick, L.K. and Davis, C.A. (2010). Seismic hazards and water supply performance. *Journal of Earthquake Engineering*, 14(7), 1022-1043.

Rose, A. and Huyck, C.K. (2016). Improving catastrophe modeling for business interruption insurance needs. *Risk Analysis*. Advance online publication. doi: 10.1111/risa.12550.

Rose, A. and Lim, D. (2002). Business interruption losses from natural hazards: conceptual and methodological issues in the case of the Northridge earthquake. *Environmental Hazards*, *4*, 1–14.

Rossetto, T. and Elnashai, A.S. (2003). Derivation of vulnerability functions for European-type RC structurs based on observational data. *Engineering Structures*, 25, 1241-1263.

Rossetto, T., Ioannou, I. and Grant, D.N. (2015). *Existing Empirical Fragility and Vulnerability Relationships: Compendium and Guide for Selection, GEM Technical Report 2015-01*. Pavia, Italy: Global Earthquake Model Foundation.

Rossetto, T., Ioannou, I., Grant, D.N. and Maqsood, T. (2014). *Guildelines for Empirical Vulnerability Assessment*. Pavia, Italy: Global Earthquake Model Foundation.

Rougier, J. (2013). Quantifying Hazard Losses. In J. Rougier, S. Sparks and L. Hill (Eds.), *Risk and Uncertainty Assessment for Natural Hazards*. Cambridge, United Kingdom: Cambridge University Press.

Ruquet, M.E. (2012). Infrastructure damage could drive Sandy losses higher than current estimates. *Property Casualty 360*. Retrieved from http://www.propertycasualty360.com/2012/11/19/infrastructure-damage-could-drive-sandy-losses-hig

Santella, N., Steinberg, L.J. and Parks, K. (2009). Decision making for extreme events: modeling critical infrastructure interdependencies to aid mitigation and response planning. *Review of Policy Research*, 26(4), 409-422.

Saunders, W., Vallance, S. and Mamula-Seadon, L. (2014). Land Use Planning Following an Earthquake Disaster. In M. Beer, I.A. Kougioumtzoglou, E. Patelli and I.S.-K. Au (Eds.), *Encyclopaedia of Earthquake Engineering*. Berlin, Germany: Springer-Verlag.

Saygili, G. and Rathje, E.M. (2008). Empirical predictive models for earthquake-induced sliding displacements of slopes. *Journal of Geotechnical and Geoenvironmental Engineering*, 134(6), 790-803.

Scherbaum, F., Cotton, F. and Smit, P. (2004). On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: the case of rock motion. *Bulletin of the Seismological Society of America*, 94(6), 2164-2185.

Schiff, A. (1997). *The Loma Prieta, California, Earthquake of October 17, 1989 – Lifelines, US Geological Survey Professional Paper 1552-A*. Boulder, CO: United States Geological Survey.

Schwartz, D.P. and Coppersmith, K.J. (1984). Fault behaviour and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones. *Journal of Geophysical Research*, 89(B7), 5681-5698.

Scordilis. E.M. (2006). Empirical global relations converting M_S and m_b to moment magnitude. *Journal of Seismology*, 10(2), 225-236.

Seed, H. and Idriss, I. (1967). Analysis of soil liquefaction: Niigata earthquake. *Journal of Soil Mechanics and Foundation Engineering*, 93(3), 83-108.

Setola, R., De Porcellinis, S. and Sforna, M. (2009). Critical infrastructure dependency assessment using the input-output inoperability model. *International Journal of Critical Infrastructure Protection*, 2, 170-178.

Shinozuka, N. and Tanaka, S. (1996). Effects of lifeline interaction under seismic condition. In *11th World Conference on Earthquake Engineering, 23-28 June 1996, Acapulco, Mexico*.

Shinozuka, M., Dong, X., Chen, T.C. and Jin, X. (2007). Seismic performance of electric transmission network under component failures. *Earthquake Engineering and Structural Dynamics*, 36, 227-244.

Shiraki, N., Shinozuka, M., Moore, J.E., Chang, S.E., Kameda, H. and Tanaka, S. (2007). System risk curves: probabilistic performance scenarios for highway networks subject to earthquake damage. *Journal of Infrastructure Systems*, 13, 43-54.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman Hall.

Sing, T., Sander, O., Beerenwinkel, N. and Lengauer, T. (2005). ROCR: visualizing classifier performance in R. *Bioinformatics*, 21(20), 3940-3941.

Sinnott, R.W. (1984). Virtues of the Haversine. *Sky and Telescope*, 68(2), 159.

Smith, T.J. and McKenna, C.M. (2013). A comparison of logistic regression psudo R^2 indices. *Multiple Linear Regression Viewpoints*, 39(2), 17-26.

Song, J. and Kang, W.-H. (2009). System reliability and sensitivity under statistical dependence by matrix-based system reliability method. *Structural Safety*, 31(2), 148-156.

Song, J. and Ok, S.-Y. (2010). Multi-scale system reliability analysis of lifeline networks under earthquake hazards. *Earthquake Engineering and Structural Dynamics*, 39, 259-279.

Statistics New Zealand. (2014a). *Census 2013*. Retrieved from http://www.stats.govt.nz/Census/2013-census.aspx

Statistics New Zealand. (2014b). *Statistical Standard for Meshblock*. http://www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/meshblock/definition.aspx

Stevenson, J.R., Kachali, H., Whitman, Z., Seville, E., Vargo, J. and Wilson, T. (2011). Preliminary observations of the impacts the 22 February Christchurch earthquake had on organisations and the economy: a report from the field (22 February 2011 – 22 March 2011). *Bulletin of the New Zealand Society for Earthquake Engineering*, 44(2), 65-76.

Steyerberg, E., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. and Kattan, M.W. (2010). Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*, 21(1), 128-138.

Stirling, M., McVerry, G. and Berryman, K. (2002). A new seismic hazard model for New Zealand. *Bulletin of the Seismological Society of America*, 92(5), 1878-1903.

Stirling, M., Gerstemberger, M., Litchfield, N., McVerry, G., Smith, W., Pettinga, J. and Barnes, P. (2008). Seismic hazard of the Canterbury region, New Zealand: new earthquake source model and methodology. *Bulletin of the New Zealand Society for Earthquake Engineering*, 41(2), 51-67.

Stirling, M., McVerry, G., Gerstenberger, M., Litchfield, N., van Dissen R., Berryman, K., Barnes, P., Wallace, L., Villamor, P., Langridge, R., Lamarche, G., Nodder, S., Reyners, M., Bradley, B.A., Rhoades, D., Smith, W., Nicol, A., Pettinga, J., Clark, K. and Jacobs, K. (2012). National seismic hazard model for New Zealand: 2010 update. *Bulletin of the Seismological Society of America*, 102(4), 1514-1542.

Stock, C. and Smith, E.G.C. (2002). Adaptive kernel estimation and continuous probability representation of historical earthquake catalogs. *Bulletin of the Seismological Society of America*, 92(3), 904-912.

Straub, D. and Der Kiureghian, A. (2008). Improved seismic fragility modeling from empirical data. *Structural Safety*, 30(4), 320-336.

Tanaka, K., Yamazaki, y., Okazawa, T., Suzuki, T., Kishimoto, T. and Iwata, K. (2008). Experiment on seismic disaster characteristics of underground cable. In *14th World Conference on Earthquake Engineering, 12-17 October 2012, Beijing, China*.

Tang, A. and Wen, A. (2009). An intelligent simulation system for earthquake disaster assessment. *Computers and Geosciences*, 35(5), 871-879.

Tatano, H. and Tsuchiya, S. (2008). A framework for economic loss estimation due to seismic transportation network disruption: a spatial computable general equilibrium approach. *Natural Hazards*, *44*, 253–265.

Tierney, K.J. (1997). Business Impacts of the Northridge Earthquake. *Journal of Contingencies and Crisis Management*, *5*(2), 87–97.

Tobin, G.A. (1999). Sustainability and community resilience: the holy grail of hazards planning? *Natural Hazards*, *1*, 13–25.

Tokimatsu, A.M. and Seed, H.B. (1987). Evaluation of settlements in sands due to earthquake shaking. *Journal of the Geotechnical Engineering Division*, 113(8), 681-878.

Toprak, S. and Taskin, F. (2006). Estimation of earthquake damage to buried pipelines caused by ground shaking. *Natural Hazards*, 40(1), 1-24.

Transpower. (2011). *22nd February 2011 Christchurch Earthquake:Key Findings and Lessons Learned*. Wellington, New Zealand: Transpower New Zealand Limited.

Ulm. K. (1990). A simple method to calculate the confidence interval of a standardized mortality ratio. *American Journal of Epidemiology*, 131(2), 373-375.

UNISDR. (2005). *Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters*. Geneva, Switzerland: United Nations Office for Disaster Reduction.

UNISDR. (2015). *Sendai Framework for Disaster Reduction 2015-2030*. Geneva, Switzerland: United Nations Office for Disaster Reduction.

USGS. (2013, January 4). *Global V_S^{30} Map Server*. Retrieved from http://earthquake.usgs.gov/hazards/apps/vs30/

USGS. (2014, October 13). *Earth Explorer*. Retrieved from http://earthexplorer.usgs.gov/

USGS. (2015a). *USGS ShakeMap: South Island of New Zealand Saturday September 4th, 2010*. Retrieved 1 March 2016 from http://earthquake.usgs.gov/earthquakes/shakemap/global/shake/2010atbj/

USGS. (2015b). *USGS ShakeMap: South Island of New Zealand Tuesday February 22nd, 2011*. Retrieved 1 March 2016 from http://earthquake.usgs.gov/earthquakes/shakemap/global/shake/b0001igm/

Van Ballegooy, S. and Malan, P. (2013). *Liquefaction Vulnerability Study, Report Prepared for the Earthquake Commission, Ref:52020.0200/v1.0*. Christchurch, New Zealand: Tonkin and Taylor.

Van Ballegooy, S., Malan, P., Lacrosse, V., Jacka, M.E., Cubrinovski, M., Bray, J.D., O'Rourke, T.D., Crawford, S.A. and Cowan, H. (2014). Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra*, 30(1), 31-56.

Vanzi, I. (1996). Seismic reliabilty of electric power networks:methodology and applicaion. *Structural Safety*, 18(4), 311-327.

Vanzi, I. (2000). Structural upgrading strategy for electric power networks under seismic action. *Earthquake Engineering and Structural Dynamics*, 29(7), 1053-1073.

Walker, G.R. (1997). Prediction of insurance loss from earthquakes. *Bulletin of the New Zealand National Society for Earthquake Engineering*, 30(1), 40-45.

Wang, M. and Takada, T. (2005). Macrospatial correlation model of seismic ground motions. *Earthquake Spectra*, 21(4), 1137-1156.

Wang, Y. (2013). Seismic risk assessment of water supply systems. In S. Tesfamariam and K.Goda (Eds.), *Handbook of Seismic risk Analysis and Management of Civil Infrastructure Systems*. Cambridge, UK: Woodhead Publishing.

Wang, Y., Au, S.-K. and Fu, Q. (2010). Seismic risk assessment and mitigation of water supply systems. *Earthquake Spectra*, 26(1), 257-274.

Wang, Y. and O'Rourke, T.D. (2008). *Seismic Performance Evaluation of Water Supply Systems, Technical Report MCEER-08-0015*, Buffalo, NY: Multidisciplinary Center for Earthquake Engineering Research (MCEER).

Weatherill, G., Esposito, S., Iervolino, I., Franchin, P. and Cavalieri, F. (2014). Framework for Seismic Hazard Analysis of Spatially Distributed Systems. In K. Pitilakis, P. Franchin, B. Khazai, H. Wenzel (Eds.), *SYNER-G: Systemic Vulnerability and Risk Assessment of Complex Urban Utility Lifeline Systems and Critical Facilities*. Dordrecht, Netherlands: Springer.

Wells, D.L. and Coppersmith, K.J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area and surface displacement. *Bulletin of the Seismological Society of America*, 84(4), 974-1002.

Wesson, R. and Perkins, D. (2001). Spatial correlation of probabilistic earthquake ground motion and loss. *Bulletin of the Seismological Society of America*, 91(6), 1498-1515.

Winkler, J., Dueñas-Osorio, L., Stein, R. and Subramanian, D. (2010). Performance assessment of topologically diverse power systems subjected to hurricane events. *Reliability Engineering and System Safety*, 95(4), 323-336.

Woo, G. (2011). *Calculating Catastrophe*. London: Imperial College Press.

Wood, C., Cox, B.R., Wotherspoon, L. and Green, R.A. (2011). Dynamic site characterization of Christchurch strong motion stations. *Bulletin of the New Zealand Society for Earthquake Engineering*, 44(4), 195-204.

Wood, P., Robins, P. and Hare, J. (2010). Preliminary observations of the 2010 Darfield (Canterbury) earthquakes: an introduction. *Bulletin of the New Zealand Society for Earthquake Engineering*, 43(4), i-iv.

Wotherspoon, L., Orense, R.P., Green, R.A., Bradley, B.A., Cox, B.R. and Wood, C. (2014). Analysis of liquefaction characteristics at Christchurch strong motion stations. In R.P. Orense, I. Towhata and N. Chouw (Eds.), *Soil Liquefaction During Recent Large-Scale Earthquakes*. Boca Raton, FL: CRC Press.

Wu, J. and Dueñas-Osorio, L. (2013). Calibration and validation of a seismic damage propagation model for interdependent infrastructure systems. *Earthquake Spectra*, 29(3), 1021-1041.

Yamada, S., Orense, R. and Cubrinovski, M. (2011). Geotechnical damage due to the 2011 Christchurch, New Zealand, earthquake. *International Society for Soil Mechanics and Geotechnical Engineering(ISSMGE) Bulletin*, 5(2), 27-45.

Youd, T. and Perkins, D. (1978). Mapping of liquefaction induced ground failure potential. *Journal of the Geotechnical Engineering Division*, 104, 433-446.

Youd, T. and Perkins, D. (1987). Mapping of liquefaction severity index. *Journal of the Geotechnical Engineering Division*, 113, 1374-1392.

Youngs, R.R., Chiou, S.-J., Silva, W.J. and Humphrey, J.R. (1997). Strong ground motion attenuation relationships for subduction zone earthquakes. *Seismological Research Letters*, 68(1), 58-73.

Zhang, J., Huo, Y., Brandenberg, S.J. and Kashighandi, P. (2008). Effects of stuctural characterizations on fragility functions of bridges subject to seismic shaking and lateral spreading. *Earthquake Engineering and Engineering Vibration*, 7(4), 369-382.

Zhou, Y.-G. and Chen, Y.-M. (2007). Laboratory investigation on assessing liquefaction resistance of sandy soils by shear wave velocity. *Journal of Geotechnical and Geoenviornmental Engineering*, 133, 959-972.

Zhu, J., Daley, D., Baise, L.G., Thompson, E.M., Wald, D.J. and Knudsen, K.L. (2015). A geospatial liquefaction model for rapid response and loss estimation. *Earthquake Spectra*, 31(3), 1818-1837.

Zimmerman, R. and Restrepo, C.E. (2009). Analyzing cascading effects within infrastructure sectors for consequence reduction. In *2009 IEEE Conference on Technologies for Homeland Security, 11-12 May 2009, Boston, MA*.

Zio, E. and Sansavini, G. (2011). Modeling interdependent network systems for identifying cascade-safe operating margins. *IEEE Transactions on Reliability*, 60(1), 94-101.

Appendix A:
List of publications

Journals

Kongar, I., Giovinazzi, S. and Rossetto, T. (2016). Seismic performance of buried electrical cables: evidence-based repair rates and fragility functions. *Bulletin of Earthquake Engineering*. Advance online publication. doi: 10.1007/s10518-016-0077-3.

Kongar, I., Rossetto, T., Giovinazzi, S. (2016). Evaluating simplified methods for liquefaction assessment for loss estimation. *Natural Hazards and Earth System Sciences*. In press. doi: 10.5194/nhess-2016-281.

Kongar, I., Esposito, S. and Giovinazzi, S. (2017). Post-earthquake assessment and management for infrastructure systems: learning from the Canterbury (New Zealand) and L'Aquila (Italy) earthquakes. *Bulletin of Earthquake Engineering*, 15, 598-620.

Books

Kongar, I. and Giovinazzi, S. (2015). Damage to Infrastructure: Modeling. In M. Beer, I.A. Kougioumtzoglou, E. Patelli, S.-K. Au, *Encyclopedia of Earthquake Engineering*. Berlin, Germany: Springer.

Giovinazzi, S., Kongar, I., Bocchini, G.M. and Ottonelli, D. (2015). Damage to Buildings: Modeling. In M. Beer, I.A. Kougioumtzoglou, E. Patelli, S.-K. Au, *Encyclopedia of Earthquake Engineering*. Berlin, Germany: Springer.

Conferences

Kongar, I. and Rossetto, T. (2012). A framework to assess the impact of seismic shocks on complex urban critical infrastructure networks. In *15th World Conference on Earthquake Engineering, 24-28 September 2012, Lisbon, Portugal*.

Kongar, I., Rossetto, T. and Giovinazzi, S. (2014). The effectiveness of existing methodologies for predicting electrical substation damage due to earthquakes in New Zealand. In *2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM2014), 13-16 July 2014, Liverpool, UK.*

Kongar, I., Rossetto, T. and Giovinazzi, S. (2014). Seismic fragility of underground electrical cables in the 2010-11 Canterbury (NZ) earthquakes. In *2nd European Conference on Earthquake Engineering and Seismology, 24-28 August 2014, Istanbul, Turkey.*

Kongar, I. and Rossetto, T. (2015). A methodological hierarchy for modelling lifelines interdependencies in risk management. In *SECED 2015 Conference: Earthquake Risk and Engineering towards a Resilient World, 9-10 July, Cambridge, UK.*

Kongar, I., Rossetto, T. and Giovinazzi, S. (2015). Evaluating desktop methods for assessing liquefaction-induced damage to infrastructure for the insurance sector. In *12th International Conference on Applications of Statistics and Probability in Civil Engineering, 12-15 July 2015, Vancouver, Canada.*

Kongar, I., Giovinazzi, S. and Rossetto, T. (2017). Seismic risk assessment of interdependent electric power and water supply systems in Christchurch, New Zealand. In *16th World Conference on Earthquake Engineering, 9-13 January 2017, Santiago, Chile.*

Appendix B: Infrastructure System Taxonomy

Electric power

Study	Components	Attributes
ATC-13 (ATC, 1985)	Generation plants	Building construction class
	Transmission lines	None
	Substations	Building construction class
ATC-25 (ATC, 1991)	Generation plants	Fossil fuel or hydro
	Transmission lines	None
	Distribution lines	None
	Transmission substations	None
	Distribution substations	None
Shinozuka and Tanaka (1996)	Gate station	None
	Transmission lines	Voltage
	Substations	Voltage
	Buried cables	Voltage
Vanzi (1996)	Generation plants	None
	Transmission lines	Voltage
	Substations	Hierarchy
American Lifelines Alliance (Taylor et al., 2005)	SCADA equipment	None
	Transmission substations	None
	Transmission lines	None
	Transmission towers	None
	Distribution substations	None
	Distribution lines	None
	Distribution transformers	None
	Office buildings	None
	Office equipment	None
RISK-UE (Pitilakis et al., 2006)	Generation plants	Capacity, Seismic design
	Substations	Hierarchy, Voltage, Seismic design
	Lines	Hierarchy, Seismic design

Study	Components	Attributes
Shinozuka et al. (2007)	Generation plants	None
	Substations	Seismic design
	Transmission lines	None
Adachi and Ellingwood (2008)	Gate stations	None
	Substations	Voltage
	Transmission lines	Voltage
	Buried cables	Voltage
Rasulo et al. (2008)	Generation plants	None
	Substations	None
	Transmission lines	Voltage
Azevedo et al. (2009)	Transformers	None
	Generation plants	Capacity, Seismic design
	Distribution circuits	Seismic design
	Substations	Voltage, Seismic design
Hazus-MH (DHS, 2009)	Generation plants	Capacity, Seismic design
	Distribution circuits	Seismic design
	Substations	Voltage, Seismic design
Reed et al. (2009)	Distribution lines	None
Kim et al. (2010)	Gate stations	Voltage, Seismic design
	Substations	Voltage, Seismic design
	Transmission lines	None
Hernandez-Fajardo and Duenas-Osorio (2011)	Gate stations	None
	Substations	Voltage
	Transmission lines	None
Leelardcharoen (2011)	Control centres	None
	Generation plants	None
	Substations	Voltage
	Transmission lines	None
Poljansek et al. (2012)	Generation plants	Capacity, Seismic design
	Substations	Voltage, Seismic design
	Transmission lines	None
Cavalieri et al. (2014)	Generation plants	Capacity, Seismic design
	Substations	None
	Transmission lines	None

Study	Components	Attributes
SYNER-G (Pitilakis et al., 2014)	Generation plants	Capacity, Seismic design
	Substations	None
	Transmission lines	None

Water supply

Study	Components	Attributes
ATC-13 (ATC, 1985)	Treatment plants	Building construction class
	Pipes	Material, Diameter, Joint, Elevation
	Pumping stations	Building construction class
	Storage tanks	Elevation
ATC-25 (ATC, 1991)	Aqueducts	None
	Pumping stations	None
	Storage reservoirs	None
	Treatment plants	None
	Storage tanks	None
	Trunk lines	None
	Wells	None
Shinozuka and Tanaka (1996)	Wells	Not described
	Pumping stations	Not described
	Storage tanks	Not described
	Pipes	Not described
Hwang et al. (1998)	Wells	None
	Pumping stations	Not described
	Storage tanks	Not described
	Pipes	Material, Diameter
American Lifelines Alliance (2001)	Aqueducts	Elevation, Redundancy, Appurtenances, Pumps
	Distribution pipes	Material, Diameter, Corrosion protection, Air valves
	Storage tanks	Foundation, Roof type, Capacity, Anchorage
	Tunnels	Construction
	Canals	Geometry, Capacity, Locations of siphons, flumes, crossings and other in-line components
	Valves and SCADA	None
Chang et al. (2002)	Wells	Not described
	Pumping stations	Not described
	Pipes	Not described

Study	Components	Attributes
RISK-UE (Pitilakis et al., 2006)	Source	Capacity, Power, Pump, Building, Seismic design
	Treatment plant	Capacity, Seismic design
	Pumping station	Capacity, Power, Building, Seismic design
	Storage tank	Capacity, Material, Size, Seismic design, Anchorage, Elevation, Roof type
	SCADA	None
	Pipes	Elevation, Material, Geometry, Joint type, Corrosiveness, Branches
	Tunnels	Construction, Ground conditions
	Canals	Construction, Reinforcement
Adachi and Ellingwood (2008)	Pumping stations	Capacity, Seismic design
	Storage tanks	Elevation, Anchorage, Material
	Pipes	Material, Joint type, Diameter, Soil conditions
Rasulo et al. (2008)	Pipes	None
Azevedo et al. (2009)	Pipes	None
	Treatment plants	Capacity, Seismic design
	Pumping stations	Capacity, Seismic design
	Wells	None
	Storage tanks	Elevation, Anchorage, Material
Hazus-MH (DHS, 2009)	Pipes	Material
	Treatment plants	Capacity, Seismic design
	Pumping stations	Capacity, Seismic design
	Wells	None
	Storage tanks	Elevation, Anchorage, Material
Javanbarg and Takada (2009)	Source	None
	Treatment plant	None
	Distribution plant	None
	Pumping stations	None
	Storage tanks	None
	Pipes	Material

Study	Components	Attributes
Tang and Wen (2009)	Wells	None
	Pumping stations	None
	Pipes	None
Kim et al. (2010)	Pumping stations	Capacity, Seismic design
	Pipes	Material
	Storage tanks	Elevation, Anchorage, Material
Wang et al. (2010)	Reservoirs	None
	Storage tanks	None
	Pumping stations	None
	Pipes	Material, Diameter
Hernandez-Fajardo and Duenas-Osorio (2011)	Pumping stations	Capacity, Seismic design
	Pipes	Material
	Storage tanks	Elevation, Anchorage, Material
SYNER-G (Pitilakis et al., 2014)	Wells	Seismic design
	Treatment plants	None
	Pumping stations	Seismic design
	Storage tanks	Material, Roof type, Anchorage, Elevation
	Canals	Reinforcement
	Tunnels	Construction, Ground conditions
	Pipes	Material

Waste water

Study	Components	Attributes
ATC-13 (ATC, 1985)	Treatment plants	Building construction class
	Pipes	Material, Diameter, Joint, Elevation
	Pumping stations	Building construction class
ATC-25 (ATC, 1991)	Mains	None
	Pumping stations	None
	Treatment plants	None
RISK-UE (Pitilakis et al., 2006)	Treatment plants	Capacity, Seismic design
	Lift station	Capacity, Seismic design
	SCADA	Anchorage, Valve intervals
	Pipes	Elevation, Material, Geometry, Joint type, Corrosiveness, Branches
	Tunnels	Construction, Ground conditions
Hazus-MH (DHS, 2009)	Pipes	Material
	Treatment plants	Capacity, Seismic design
	Lift stations	Capacity, Seismic design
SYNER-G (Pitilakis et al., 2014)	Treatment plants	Seismic design
	Lift stations	Seismic design
	Tunnels	Construction, Ground conditions
	Pipes	Material

Natural gas

Study	Components	Attributes
ATC-13 (ATC, 1985)	Holding facilities	Building construction class
	Compressor stations	Building construction class
	Transmission lines	Material, Diameter, Joint, Elevation
	Distribution mains	Material, Diameter, Joint, Elevation
ATC-25 (ATC, 1991)	Transmission lines	None
	Compressor stations	None
	Distribution mains	None
RISK-UE (Pitilakis et al., 2006)	Production plants	Seismic design
	Tank farms	Seismic design
	Pipelines	Material, Diameter, Joint type, Soil conditions
	Stations	Seismic design
Chang and Song (2007)	Stations	None
	Pipelines	Soil conditions, Material, Joint type
Azevedo et al. (2009)	Pipelines	Material
	Compressor stations	Seismic design
Hazus-MH (DHS, 2009)	Pipelines	Material
	Compressor stations	Seismic design
Poljansek et al. (2012)	Pipelines	Material
	Compressor stations	Seismic design
SYNER-G (Pitilakis et al., 2014)	Pipelines	Material, Diameter, Joint type, Soil conditions
	Compressor stations	Seismic design
	Tank farms	Seismic design
Esposito et al. (2015)	Pipelines	Material, Diameter, Joint type, Soil conditions
	Compressor stations	Seismic design
	Tank farms	Seismic design
	Regulators	None

Oil delivery

Study	Components	Attributes
ATC-25 (ATC, 1991)	Oil fields	None
	Refineries	None
	Transmission pipelines	None
	Storage tanks	None
Azevedo et al. (2009)	Refineries	Capacity, Seismic design
	Pipelines	Material
	Pumping stations	Seismic design
	Tank farms	Seismic design
Hazus-MH (DHS, 2009)	Refineries	Capacity, Seismic design
	Pipelines	Material
	Pumping stations	Seismic design
	Tank farms	Seismic design
SYNER-G (Pitilakis et al., 2014)	Processing plants	
	Pipelines	Material, Diameter, Joint type, Soil conditions
	Pumping stations	Seismic design
	Tank farms	Seismic design

Telecoms

Study	Components	Attributes
ATC-13 (ATC, 1985)	Switching offices	Building construction class
	Towers	None
RISK-UE (Pitilakis et al., 2006)	Central offices	Anchorage
	Cables	Material, Installation, Depth
Hazus-MH (DHS, 2009)	Central offices	Seismic design
	Transmission lines	None
	Cabling	None
Reed et al., 2009	Lines	None
Leelardcharoen (2011)	Lines	None
	Offices	None

Highways

Study	Components	Attributes
ATC-13 (ATC, 1985)	Bridges	Hierarchy
	Tunnels	None
	Roads	Hierarchy
	Terminal stations	Building construction class
ATC-25 (ATC, 1991)	Bridges	None
	Tunnels	None
	Roads	None
Kawakami (2000)	Roads	Hierarchy
RISK-UE (Pitilakis et al., 2006)	Roads	Geometry, Topography, Building environment
	Bridges	Material, Structure, Pier type, Connection type, Bearing type, Span, Age, Number of spans, Skew Span length, Height, Number of expansion joints, Foundation type
	Tunnels	Construction, Ground conditions
Shiraki et al. (2007)	Bridges	None
	Roads	None
Rasulo et al. (2008)	Bridges	Not described
	Tunnels	Not described
	Roads	Not described
	Embankments	Not described
	Retaining walls	Not described
Azevedo et al. (2009)	Roads	None
	Bridges	Material, Number of spans, Connection type, Deck type, Number of column alignments, Environment, Age
	Viaducts	same as for bridges
	Tunnels	Not described

Study	Components	Attributes
Hazus-MH (DHS, 2009)	Roads	Hierarchy
	Bridges	Age, Construction, Seismic design, Number of spans, Skew, Width, Length, Maximum span
	Tunnels	Construction
Guikema and Gardoni (2009)	Bridges	Span, Centre span, Geometry, Soil stiffness, Material, Dead loads
Jayaram and Baker (2010)	Bridges	Age, Construction, Seismic design, Number of spans, Skew, Width, Length, Maximum span
SYNER-G (Pitilakis et al., 2014)	Tunnels	Construction, Ground conditions, Hierarchy
	Embankments	Height
	Trenches	None
	Slopes	None
	Roads	Number of lanes
	Bridge abutments	Height

Appendix C: Infrastructure Fragility Functions

Electric power

Study	IM(s)	Typologies
Generation plants		
Hazus (DHS, 2009)	PGA, PGDf	Capacity, Seismic design
Substations		
Ang et al. (1996)	PGA	Voltage
Hwang and Chou (1998)	PGA	Seismic design
Vanzi (2000)	PGA	None
Rasulo et al. (2004)	PGA	None
Hazus (DHS, 2009)	PGA, PGDf	Voltage, Seismic design
Distribution circuits		
Hazus (DHS, 2009)	PGA, PGDf	Seismic design
Conduits		
Park et al. (2006)	PGA, PGV, MMI	None
Circuit breakers		
Agnanos and Ostrom (2000)	PGA	None
Rasulo et al. (2004)	PGA	None
Vanzi et al. (2004)	PGA	None
Shinozuka et al. (2007)	PGA	None
Straub and Der Kiureghian (2008)	PGA	None
Lightning arresters		
Hwang and Huo (1998)	PGA	None
Agnanos (1999)	PGA	Voltage
Disconnect switch		
Ang et al. (1996)	PGA	None
Agnanos (1999)	PGA	Voltage, Seismic design
Agnanos and Ostrom (2000)	PGA	None
Shinozuka et al. (2007)	PGA	None

Study	IM(s)	Typologies
Transformers		
Hwang and Huo (1998)	PGA	None
Agnanos (1999)	PGA	None
Liu et al. (2003)	PGA	Voltage
Shinozuka et al. (2007)	PGA	None
Straub and Der Kiureghian (2008)	PGA	None
Coil support		
Rasulo et al. (2004)	PGA	None
Current transformer		
Rasulo et al. (2004)	PGA	None
Voltage transformer		
Rasulo et al. (2004)	PGA	None
Horizontal sectionalising switch		
Rasulo et al. (2004)	PGA	None
Vertical sectionalising switch		
Rasulo et al. (2004)	PGA	None
Discharger		
Rasulo et al. (2004)	PGA	None
Bar support		
Rasulo et al. (2004)	PGA	None
Autotransfomer		
Rasulo et al. (2004)	PGA	None
Box		
Rasulo et al. (2004)	PGA	None
Power supply		
Rasulo et al. (2004)	PGA	None
Pothead		
Hwang and Huo (1998)	PGA	None
Control house		
Hwang and Huo (1998)	PGA	None
Capacitor bank		
Hwang and Huo (1998)	PGA	None

Study	IM(s)	Typologies
Regulator		
Hwang and Huo (1998)	PGA	None
Bus		
Shinozuka et al. (2007)	PGA	None

Water supply and Waste water

Study	IM(s)	Typologies
Pipes		
Katayama et al. (1975)	PGA	Soil conditions, Age
Isoyama and Katayama (1982)	PGA	Cast iron only
Barenberg (1988)	PGV	Material
Hamada (1991)	PGA	None
Porter et al. (1991)	PGDf	Material
Wang et al. (1991)	MSK intensity	Soil conditions
Honegger and Eguchi (1993)	PGDf	Material
O'Rourke and Ayala (1993)	PGV	Material
Heubach (1995)	PGDf	Material, Joint type
Ballantyne and Heubach (1996)	PGDf	None
Eidinger (1998)	PGV	Material
Isoyama (1998)	PGV	Material, Diameter
O'Rourke et al. (1998)	PGA	None
Toprak (1998)	PGA	None
Eidinger et al. (1999)	PGDf	Material, Joint type
O'Rourke and Leon (1999)	PGV	Diameter
Eidinger and Avila (1999)	PGV	Material, Diameter, Soil conditions, Joint type
Isoyama et al. (2000)	PGA	Cast iron only
ALA (2001	PGV, PGDf	Material, Joint type
Hung (2001)	PGA	None
Pineda-Parros and Ordaz (2003)	PGV	None
O'Rourke and Deyoe (2004)	PGV	Ductility, Buried or above ground
Pineda-Parros and Ordaz (2007)	PGV^2/PGA	None
Yeh et al. (2006)	PGA	None
O'Rourke et al. (2014)	PGV, PGDf	Material
Sources		
Hazus (DHS, 2009)	PGA, PGDf	Seismic design
Treatment plants		
Hazus (DHS, 2009)	PGA, PGDf	Capacity, Seismic design
SRMLIFE (2007)	PGA	Building type, Seismic design
Pumping/Lift stations		
Hazus (DHS, 2009)	PGA, PGDf	Seismic design
SRMLIFE (2007)	PGA	Seismic design

Study	IM(s)	Typologies
Storage tanks		
O'Rourke and So (2000)	PGA	Height/diameter ratio, Quantity stored content
Hazus (DHS, 2009)	PGA	Material, Elevation
ALA (2001)	PGA, PGDf	Anchorage, Material, Size, Seismic design, Roof
Tunnels		
ALA (2001)	PGA	Ground conditions
Salmon et al. (2003)	PGA, PGDf	Construction method
Argyroudis and Pitilakis (2007)	PGDf	Ground conditions, Construction method
Corigliano (2007)	PGV	None
Hazus (DHS, 2009)	PGA, PGDf	Construction method
Argyroudis (2010)	PGA	Ground conditions, Construction method
Canals		
ALA (2001)	PGV, PGDf	None

Natural gas and Oil delivery

Study	IM(s)	Typologies
Pumping/Compressor station		
LESSLOSS (2007)	PGA	Seismic design
Hazus (DHS, 2009)	PGA, PGDf	Capacity, Seismic design
SRMLIFE (2007)	PGA	Building type, Seismic design
Storage tanks		
O'Rourke and So (2000)	PGA	Height/diameter ratio, Quantity stored content
ALA (2001)	PGA, PGDf	Anchorage, Material, Size, Seismic design, Roof
Iervolino et al. (2004)	PGA	Height/diameter ratio, Friction at base, Steel only
Berahman and Behnamfar (2007)	PGA	Anchorage, Elevation, Steel only
Hazus (DHS, 2009)	PGA, PGDf	Material, Elevation
Pipelines		
Same as pipes in water supply system		

Telecoms

Study	IM(s)	Typologies
Central office		
Hazus (DHS, 2009)	PGA, PGDf	Seismic design
Point of presence		
Leelardcharoen (2011)	PGA	None
Tandem office		
Leelardcharoen (2011)	PGA	None
End office		
Leelardcharoen (2011)	PGA	None

Highways

Study	IM(s)	Typologies
Tunnels		
Same as water supply system		
Embankments		
Lagaros et al. (2009)	PGA	None
Maruyama et al. (2010)	PGV	None
Bridges		
Gardoni et al. (2003)	SA	Two-span concrete only
Choi et al. (2004)	PGA	Material, Span type
Elnashai et al. (2004)	PGA	Material, Pier diameter, Centre span, External span
Mackie and Stojadinovic (2004)	SA	Two-span concrete only
Karim and Yamazaki (2001)	PGA, PGV	RC only
Nielson and DesRoches (2007)	PGA	Multi-span supported concrete girder only
Banerjee and Shinozuka (2008)	PGA	RC only
Padgett and DesRoches (2008)	PGA	Multi-span continuous concrete girder only
Zhang et al. (2008)	PGA, PGDf	Deck design, Colum type, Bearing type, Gap distance
Choe et al. (2009)	SA	RC only
Gardoni and Rosowsky (2009)	SA	RC only
Hazus (DHS, 2009)	SA, PGDf	NBI class, Age, State, No. of spans, Max. span length,
Moschonas et al. (2009)	PGA	Length, Seismic design,
Kwon and Elnashai (2010)	PGA	Pier type, Deck type, Pier-to-deck connection type Multi-span continuous steel girder only
Slopes		
ATC-13 (1985)	MMI	Slope class
Pitilakis et al. (2010)	PGA	Hierarchy level, Slope yield coefficient
Roads		
Hazus (DHS, 2009)	PGDf	Hierarchy level
Retaining walls		
ATC-13 (1985)	MMI	None
Salmon et al. (2003)	PGA	None

Appendix D:
Christchurch Pipe Fragility Analysis

Zone A – No liquefaction

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
AC	PGV	0.548	0.022	0.713	0.004
	$PGDf_H$	0.465	0.01	0.266	0.072
	$PGDf_V$	0.896	$3.36x10^{-5}$	0.739	0.001
	$PGDf_{MAX}$	0.653	0.001	0.776	0.0002
	$PGDf_{VECT}$	0.829	$6.1x10^{-6}$	0.845	$3.41x10^{-6}$
	$PGDf_{GEOM}$	0.829	$3.9x10^{-5}$	0.817	$5.48x10^{-5}$
CI	PGV	0.358	0.156	0.261	0.241
	$PGDf_H$	0.105	0.479	0.224	0.283
	$PGDf_V$	0.311	0.329	0.064	0.681
	$PGDf_{MAX}$	0.849	0.009	0.855	0.008
	$PGDf_{VECT}$	0.537	0.061	0.539	0.06
	$PGDf_{GEOM}$	0.665	0.048	0.529	0.101
GALV	PGV	0.533	0.062	0.399	0.128
	$PGDf_H$	0.385	0.024	0.132	0.223
	$PGDf_V$	0.728	0.0008	0.581	0.006
	$PGDf_{MAX}$	0.453	0.008	0.476	0.006
	$PGDf_{VECT}$	0.565	0.002	0.63	0.0007
	$PGDf_{GEOM}$	0.527	0.011	0.571	0.007
HDPE	PGV	0.83	0.0006	0.9	$9.7x10^{-5}$
	$PGDf_H$	0.463	0.015	0.11	0.293
	$PGDf_V$	0.775	0.0002	0.596	0.003
	$PGDf_{MAX}$	0.874	$1.12x10^{-7}$	0.71	$4.18x10^{-5}$

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
	$PGDf_{VECT}$	0.603	0.001	0.399	0.015
	$PGDf_{GEOM}$	0.869	2.96E-05	0.704	0.001
MDPE80	PGV	0.813	0.036	0.809	0.038
	$PGDf_V$	0.072	0.661	0.117	0.573
Brittle	PGV	0.539	0.016	0.746	0.001
	$PGDf_H$	0.707	1.25E-05	0.538	0.0005
	$PGDf_V$	0.416	0.001	0.157	0.068
	$PGDf_{MAX}$	0.627	1.13E-05	0.716	7.01E-07
	$PGDf_{VECT}$	0.758	3.2E-08	0.842	2.78E-10
	$PGDf_{GEOM}$	0.479	0.003	0.713	3.91E-05
Ductile	PGV	0.83	0.0006	0.9	9.7E-05
	$PGDf_H$	0.463	0.015	0.11	0.293
	$PGDf_V$	0.775	0.0002	0.596	0.003
	$PGDf_{MAX}$	0.874	1.12E-07	0.71	4.18E-05
	$PGDf_{VECT}$	0.603	0.001	0.399	0.015
	$PGDf_{GEOM}$	0.869	2.96E-05	0.704	0.001

Zone B – All liquefaction

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
AC	PGV	0.872	0.0007	0.788	0.003
	$PGDf_H$	0.503	0.001	0.598	0.0002
	$PGDf_V$	0.184	0.217	0.085	0.415
	$PGDf_{MAX}$	0.404	0.006	0.382	0.008
	$PGDf_{VECT}$	0.512	0.001	0.725	$1.44x10^{-5}$
	$PGDf_{GEOM}$	0.65	0.003	0.807	0.0002
CI	PGV	0.003	0.925	0.001	0.965
	$PGDf_H$	0.422	0.03	0.392	0.039
	$PGDf_V$	0.197	0.319	0.173	0.353
	$PGDf_{MAX}$	0.639	0.006	0.689	0.003
	$PGDf_{VECT}$	0.44	0.037	0.576	0.011
	$PGDf_{GEOM}$	0.897	0.001	0.893	0.001
CLS	$PGDf_H$	0.777	0.048	0.747	0.059
	$PGDf_{MAX}$	0.6	0.07	0.423	0.162
	$PGDf_{VECT}$	0.651	0.052	0.558	0.088
	$PGDf_{GEOM}$	0.846	0.08	0.888	0.057
GALV	PGV	0.484	0.055	0.45	0.068
	$PGDf_H$	0.089	0.244	0.027	0.526
	$PGDf_V$	0.174	0.303	0.264	0.193
	$PGDf_{MAX}$	0.135	0.178	0.051	0.417
	$PGDf_{VECT}$	0.014	0.674	0.001	0.927
	$PGDf_{GEOM}$	0.129	0.309	0.018	0.709
HDPE	PGV	0.087	0.479	0.069	0.53
	$PGDf_H$	0.006	0.789	0.056	0.416
	$PGDf_V$	0.174	0.264	0.049	0.569
	$PGDf_{MAX}$	0	0.956	0.031	0.567

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
	$PGDf_{VECT}$	0.058	0.406	0.001	0.92
	$PGDf_{GEOM}$	0.334	0.08	0.098	0.379
MDPE80	PGV	0.042	0.696	0.075	0.599
	$PGDf_{H}$	0.91	$1.9x10^{-5}$	0.806	0.0004
	$PGDf_{V}$	0.341	0.223	0.133	0.478
	$PGDf_{MAX}$	0.766	0.002	0.413	0.062
	$PGDf_{VECT}$	0.859	$4.06x10^{-5}$	0.884	$1.65x10^{-5}$
	$PGDf_{GEOM}$	0.826	0.002	0.825	0.002
PVC	PGV	0.539	0.266	0.58	0.239
	$PGDf_{H}$	0.771	0.021	0.569	0.083
	$PGDf_{V}$	0.424	0.234	0.246	0.395
	$PGDf_{MAX}$	0.725	0.067	0.752	0.057
	$PGDf_{VECT}$	0.393	0.183	0.468	0.134
	$PGDf_{GEOM}$	0.667	0.047	0.708	0.036
STEEL	$PGDf_{MAX}$	0.809	0.1	0.753	0.132
	$PGDf_{GEOM}$	0.086	0.707	0.106	0.674
Brittle	PGV	0.864	0.0003	0.772	0.002
	$PGDf_{H}$	0.58	$2.45x10^{-5}$	0.571	$3.03x10^{-5}$
	$PGDf_{V}$	0.015	0.679	0.004	0.833
	$PGDf_{MAX}$	0.576	$4.22x10^{-5}$	0.569	$5.05x10^{-5}$
	$PGDf_{VECT}$	0.551	$2.13x10^{-5}$	0.411	0.0006
	$PGDf_{GEOM}$	0.506	0.004	0.342	0.028
Ductile	PGV	0.001	0.946	0.023	0.698
	$PGDf_{H}$	0.477	0.002	0.164	0.096
	$PGDf_{V}$	0.153	0.297	0.125	0.999
	$PGDf_{MAX}$	0.66	$4.16x10^{-5}$	0.425	0.003
	$PGDf_{VECT}$	0.563	0.0005	0.357	0.011

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
	$PGDf_{GEOM}$	0.674	0.002	0.338	0.06

Zone C – Liquefaction, settlement only

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
AC	PGV	0.847	0.001	0.825	0.002
	$PGDf_H$	0.568	0.0007	0.643	0.0002
	$PGDf_V$	0.212	0.213	0.088	0.44
	$PGDf_{MAX}$	0.667	0.0001	0.611	0.0003
	$PGDf_{VECT}$	0.492	0.002	0.673	$9.88x10^{-5}$
	$PGDf_{GEOM}$	0.817	0.0008	0.877	0.0002
CI	PGV	0.358	0.286	0.374	0.273
	$PGDf_H$	0.341	0.076	0.23	0.16
	$PGDf_V$	0.841	0.028	0.702	0.171
	$PGDf_{MAX}$	0.683	0.003	0.667	0.004
	$PGDf_{VECT}$	0.563	0.02	0.556	0.021
	$PGDf_{GEOM}$	0.755	0.011	0.732	0.014
GALV	PGV	0.748	0.006	0.608	0.023
	$PGDf_H$	0.18	0.089	0.061	0.339
	$PGDf_V$	0.073	0.556	0.134	0.419
	$PGDf_{MAX}$	0.252	0.057	0.16	0.139
	$PGDf_{VECT}$	0.082	0.301	0.026	0.564
	$PGDf_{GEOM}$	0.004	0.873	0.013	0.77
HDPE	PGV	0.058	0.567	0.064	0.546
	$PGDf_H$	0.271	0.082	0.27	0.083
	$PGDf_V$	0.054	0.656	0.073	0.604
	$PGDf_{MAX}$	0.087	0.407	0.123	0.319
	$PGDf_{VECT}$	0.342	0.034	0.367	0.028
	$PGDf_{GEOM}$	0.119	0.403	0.212	0.251
MDPE80	PGV	0.005	0.889	0.007	0.878
	$PGDf_H$	0	0.999	0.014	0.882

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
	$PGDf_V$	0.276	0.363	0.273	0.367
	$PGDf_{GEOM}$	0.548	0.152	0.364	0.282
PVC	$PGDf_H$	0.052	0.771	0.064	0.747
	$PGDf_{MAX}$	0.781	0.116	0.759	0.129
	$PGDf_{GEOM}$	0.626	0.209	0.554	0.256
Brittle	PGV	0.927	$3.13x10^{-5}$	0.909	$6.88x10^{-5}$
	$PGDf_H$	0.503	0.0007	0.493	0.0008
	$PGDf_V$	0.379	0.025	0.239	0.09
	$PGDf_{MAX}$	0.368	0.006	0.45	0.002
	$PGDf_{VECT}$	0.428	0.002	0.344	0.007
	$PGDf_{GEOM}$	0.488	0.011	0.196	0.149
Ductile	PGV	0.068	0.496	0.086	0.443
	$PGDf_H$	0.003	0.853	0.054	0.404
	$PGDf_V$	0.272	0.185	0.182	0.292
	$PGDf_{MAX}$	0.096	0.303	0.326	0.042
	$PGDf_{VECT}$	0.15	0.154	0.009	0.731
	$PGDf_{GEOM}$	0.078	0.435	0	0.988

Zone D – Liquefaction, with lateral spread

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
AC	PGV	0.073	0.605	0.073	0.604
	$PGDf_H$	0.726	0.002	0.75	0.001
	$PGDf_V$	0.108	0.524	0.125	0.492
	$PGDf_{MAX}$	0.397	0.069	0.334	0.103
	$PGDf_{VECT}$	0.357	0.089	0.224	0.198
	$PGDf_{GEOM}$	0.047	0.641	0.015	0.272
GALV	$PGDf_V$	0.984	0.008	0.982	0.009
	$PGDf_{GEOM}$	0.31	0.33	0.331	0.31
HDPE	PGV	0.001	0.966	0.001	0.955
	$PGDf_H$	0.001	0.949	0.015	0.774
	$PGDf_V$	0.024	0.803	0.058	0.695
	$PGDf_{MAX}$	0.171	0.356	0.095	0.502
	$PGDf_{VECT}$	0.223	0.345	0.256	0.305
	$PGDf_{GEOM}$	0.408	0.122	0.299	0.204
MDPE80	PGV	0.045	0.733	0.057	0.7
	$PGDf_V$	0.948	0.026	0.97	0.015
	$PGDf_{GEOM}$	0.842	0.028	0.844	0.027
Brittle	PGV	0.033	0.731	0.042	0.698
	$PGDf_H$	0	0.967	0.007	0.772
	$PGDf_V$	0.312	0.193	0.366	0.15
	$PGDf_{MAX}$	0.001	0.794	0.048	0.473
	$PGDf_{VECT}$	0.02	0.63	0.059	0.403
	$PGDf_{GEOM}$	0.214	0.21	0.115	0.373
Ductile	PGV	0.859	0.023	0.881	0.018
	$PGDf_H$	0.139	0.233	0.027	0.608
	$PGDf_V$	0.03	0.71	0.075	0.551

Material	IM	Linear form		Logarithmic form	
		R^2	p-value	R^2	p-value
	$PGDf_{MAX}$	0.578	0.007	0.504	0.014
	$PGDf_{VECT}$	0.557	0.008	0.63	0.004
	$PGDf_{GEOM}$	0.376	0.106	0.433	0.076

Appendix E:
Interdependency event summary

This appendix summarises the key data from the events in the simulated seismic risk assessment catalogue that are relevant to the simulation and quantification of interdependency effects in the water supply system. This is defined as the set of events in which there some loss of performance is observed in the electric power network and hence there is potential for the effect of interdependency to be activated. There are 228 events in which this criterion is observed and their source locations are shown in Figure E1. The subsequent tables summarise the characteristics of these events and the corresponding engineering performance.

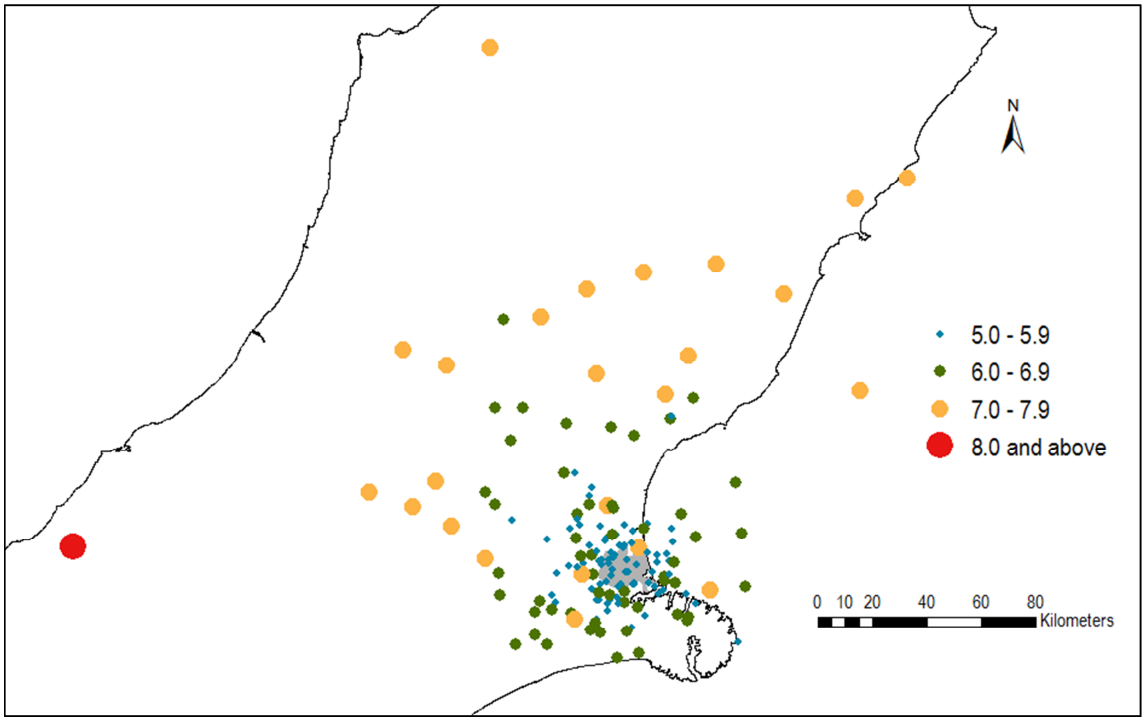

Figure E1 – Source locations and magnitudes of 228 events in seismic risk assessment catalogue with interdependency potential due to loss of performance in the electric power network

Event characteristics

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
36	7.1	-42.953	172.814	10	66	0.26	46	141	48
74	6.3	-43.209	172.349	20	40	0.29	53	115	39
164	7.0	-43.194	173.775	17	103	0.28	29	125	43
170	5.6	-43.473	172.690	15	10	0.48	61	140	48
236	5.6	-43.433	172.538	12	10	0.67	60	88	30
320	7.0	-43.459	172.687	19	11	1.11	182	231	79
355	7.6	-42.936	173.691	11	111	0.25	40	111	38
459	5.7	-43.505	172.828	15	20	0.78	103	102	35
479	6.9	-43.045	172.366	10	56	0.22	42	39	13
633	5.5	-43.440	172.842	10	23	0.57	42	74	25
869	5.3	-43.287	172.467	14	28	0.85	46	72	24
1085	7.4	-43.317	171.663	11	78	0.28	40	109	37
1134	6.4	-43.086	172.669	13	49	0.38	38	102	35
1202	6.7	-43.810	172.687	14	33	0.63	70	197	67
1213	5.7	-43.420	172.321	14	199	0.50	41	33	11
1263	5.6	-43.532	172.570	1	2	1.39	210	179	61
1304	6.9	-42.964	172.936	10	68	0.21	27	41	14
1426	8.1	-42.288	170.115	11	243	0.19	42	66	22
1428	6.6	-43.018	172.394	37	58	0.19	35	21	7
1468	6.2	-43.483	172.475	0	10	1.02	155	194	66
1516	7.6	-42.936	173.691	11	111	0.21	50	79	27
1537	6.9	-43.101	172.111	13	60	0.44	51	195	66
1603	5.7	-43.509	172.850	9	22	0.97	69	123	42
1611	5.4	-43.575	172.665	10	9	0.73	43	49	17
1628	5.0	-43.516	172.678	1	8	0.65	98	53	18

[2] Distance is from a reference location of the Department of Civil and Natural Resources Building at the University of Canterbury in Christchurch

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
1636	5.8	-43.581	172.661	19	9	0.37	49	44	15
1645	5.4	-43.724	172.655	8	23	0.46	44	13	4
1672	6.5	-43.427	172.407	19	18	0.64	136	183	62
1721	6.5	-43.639	172.622	11	13	0.67	106	154	52
1789	5.7	-43.694	172.498	6	20	1.08	110	126	43
1797	7.6	-42.936	173.691	11	111	0.27	38	118	40
1802	7.1	-42.546	172.716	10	109	0.31	47	110	37
1835	7.2	-43.380	171.838	10	62	0.28	51	91	31
1837	7.5	-42.230	173.891	10	179	0.24	49	117	40
1935	5.8	-43.712	172.465	5	23	0.70	63	119	40
1953	7.1	-42.881	172.503	10	71	0.20	42	48	16
1987	5.8	-43.397	172.838	13	25	0.78	129	141	48
2213	5.3	-43.484	172.581	5	4	1.68	253	156	53
2241	5.7	-43.210	172.400	33	38	0.40	66	35	12
2245	7.7	-42.519	173.043	11	118	0.16	62	80	27
2248	5.8	-43.371	172.697	37	19	0.36	63	37	13
2279	6.1	-43.311	172.037	8	50	0.44	27	70	24
2281	6.9	-43.045	172.366	10	56	0.28	38	101	34
2285	6.6	-43.414	173.155	15	48	0.47	65	178	61
2291	5.6	-43.771	173.141	16	53	0.31	23	13	4
2320	6.8	-43.315	172.467	10	25	0.63	92	198	67
2428	5.3	-43.504	172.509	5	6	1.15	103	138	47
2507	5.9	-43.396	172.375	9	22	0.41	52	102	35
2651	5.5	-43.633	172.558	18	13	0.51	36	40	14
2733	5.2	-43.451	172.622	8	8	0.82	67	63	21
2782	6.0	-43.602	172.625	4	10	1.44	186	213	72
2785	5.4	-43.381	172.728	7	19	0.67	109	164	56
2807	5.6	-43.611	172.779	17	19	0.37	51	38	13

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
2886	5.7	-43.612	172.902	14	28	0.54	71	102	35
2916	7.6	-42.936	173.691	11	111	0.36	72	211	72
3029	7.7	-42.519	173.043	11	118	0.33	78	147	50
3169	6.2	-43.246	173.131	9	54	0.33	28	19	6
3233	6.8	-43.008	172.011	14	74	0.37	51	112	38
3527	7.5	-43.317	172.547	11	23	0.90	141	233	79
3688	6.3	-43.349	171.917	11	57	0.34	31	61	21
3746	5.1	-43.463	172.682	14	10	0.47	40	30	10
3754	5.3	-43.477	172.561	7	5	0.61	76	133	45
3805	7.2	-42.823	172.917	10	82	0.34	66	161	55
3809	5.8	-43.366	172.114	7	42	0.55	35	124	42
3848	6.6	-43.030	172.834	10	58	0.44	45	115	39
3855	7.4	-42.602	172.459	11	103	0.16	33	41	14
3871	5.5	-43.452	172.604	4	8	0.84	118	128	44
3932	5.8	-43.645	172.947	20	32	0.43	47	12	4
3941	5.4	-43.600	172.799	16	19	0.58	39	83	28
3997	5.8	-43.419	172.650	3	13	0.92	133	130	44
4086	7.3	-42.793	171.632	11	112	0.23	51	47	16
4216	7.5	-43.317	172.547	11	23	0.46	126	215	73
4225	6.1	-43.677	172.378	2	24	0.73	148	171	58
4353	6.2	-43.576	172.853	12	23	1.03	85	194	66
4360	5.8	-43.108	172.679	10	47	0.47	57	45	15
4373	6.1	-43.425	172.949	3	31	0.40	82	95	32
4428	5.5	-43.666	172.545	5	16	1.29	139	108	37
4433	5.8	-43.258	172.477	15	31	0.35	29	47	16
4492	6.4	-43.682	172.866	11	29	0.68	82	128	44
4496	6.6	-43.617	172.559	2	11	1.11	418	230	78
4563	6.1	-43.634	172.235	15	31	0.59	71	96	33

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
4678	6.0	-43.547	172.479	10	9	0.80	92	164	56
4741	6.2	-43.613	172.054	20	44	0.57	67	43	15
4823	5.4	-43.580	172.667	3	9	1.62	223	154	52
4904	5.3	-43.618	172.742	6	17	0.46	39	9	3
4944	6.1	-43.394	172.711	13	17	1.26	242	222	76
4948	5.6	-43.538	172.630	7	4	1.40	225	199	68
4969	6.0	-43.554	172.806	9	18	0.60	93	127	43
4997	7.1	-42.848	171.828	10	97	0.28	47	123	42
5030	7.7	-41.797	172.037	11	197	0.14	32	21	7
5033	6.7	-43.739	172.512	4	25	0.55	134	197	67
5055	5.5	-43.613	172.290	6	26	0.60	48	48	16
5136	5.9	-43.025	172.842	5	59	0.31	19	6	2
5240	7.5	-42.302	173.664	10	162	0.32	57	179	61
5293	6.0	-43.690	172.914	16	33	0.18	26	2	1
5308	6.6	-43.443	172.342	2	21	0.62	83	153	52
5413	5.2	-43.581	172.607	8	7	0.78	86	47	16
5581	5.8	-43.382	172.623	15	16	0.74	72	119	40
5687	7.3	-42.620	173.343	10	118	0.20	37	62	21
5767	5.6	-43.416	172.585	12	12	0.64	52	72	24
5807	7.4	-43.317	171.663	11	78	0.35	51	117	40
5816	5.6	-43.583	172.600	11	7	0.59	110	130	44
6107	7.5	-42.302	173.664	10	162	0.24	42	145	49
6273	5.5	-43.442	172.659	4	11	1.21	111	142	48
6305	5.5	-43.588	172.597	13	7	0.76	60	120	41
6385	7.4	-43.317	171.663	11	78	0.29	46	162	55
6470	7.3	-42.793	171.632	11	112	0.25	43	52	18
6519	5.3	-43.541	172.642	12	5	0.83	69	87	30
6538	6.1	-43.669	172.214	13	34	0.45	40	55	19

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
6547	7.1	-42.546	172.716	10	109	0.14	26	1	0
6635	5.8	-43.387	172.671	11	17	0.45	88	95	32
6739	6.5	-43.345	172.409	14	24	0.36	65	148	50
6806	5.4	-43.364	172.408	3	22	0.49	35	25	9
6817	5.4	-43.400	172.451	6	17	0.27	36	2	1
6818	5.1	-43.481	172.813	16	19	0.53	38	61	21
6865	5.7	-43.430	172.274	2	27	0.33	38	15	5
7024	6.2	-43.567	172.805	14	19	1.45	108	241	82
7056	7.5	-43.317	172.547	11	23	0.51	93	223	76
7107	5.9	-43.683	172.717	10	21	0.87	45	83	28
7132	5.8	-43.484	172.771	2	16	1.37	266	194	66
7165	6.9	-43.045	172.366	10	56	0.52	44	128	44
7189	7.7	-42.056	173.656	11	185	0.14	42	22	7
7227	5.2	-43.645	172.594	10	14	0.83	52	93	32
7269	6.2	-43.702	172.910	17	33	1.38	68	123	42
7287	5.1	-43.581	172.512	1	9	1.38	113	112	38
7294	6.5	-44.336	172.678	24	91	0.37	38	46	16
7331	5.5	-43.595	172.475	5	12	0.34	67	20	7
7347	5.8	-43.632	172.509	7	14	0.66	108	145	49
7506	5.1	-43.501	172.637	0	5	2.15	245	183	62
7514	6.5	-43.271	171.993	14	55	0.22	35	26	9
7567	7.7	-42.519	173.043	11	118	0.24	42	118	40
7589	5.5	-43.576	172.735	17	14	0.75	78	40	14
7688	6.7	-44.114	171.789	36	92	0.19	40	27	9
7785	7.4	-43.317	171.663	11	78	0.37	42	151	51
7873	6.4	-43.349	172.882	12	31	0.98	112	174	59
7902	7.4	-43.317	171.663	11	78	0.23	45	107	36
8045	5.8	-43.638	172.309	13	26	0.74	62	157	53

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
8187	7.6	-42.936	173.691	11	111	0.29	45	121	41
8384	5.6	-43.418	172.467	11	15	0.97	92	102	35
8445	5.6	-43.497	172.519	14	6	0.28	36	9	3
8458	5.4	-43.644	172.543	5	14	0.60	62	38	13
8524	6.1	-43.588	173.174	6	48	0.39	37	31	11
8552	5.0	-43.591	172.718	5	13	0.57	48	16	5
8716	7.6	-42.936	173.691	11	111	0.38	80	123	42
8839	5.0	-43.552	172.564	4	4	0.78	104	90	31
8983	5.3	-43.512	172.436	11	12	0.49	42	55	19
9158	5.2	-43.510	172.774	10	15	0.37	42	23	8
9188	7.3	-42.793	171.632	11	112	0.17	28	46	16
9196	7.5	-43.317	172.547	11	23	0.86	125	223	76
9239	5.2	-43.540	172.750	6	14	1.29	113	60	20
9326	5.6	-43.309	172.526	19	24	0.81	48	86	29
9404	5.6	-43.458	172.618	4	8	0.88	185	157	53
9408	7.0	-43.697	172.396	4	25	2.50	247	254	86
9495	6.4	-43.777	172.267	6	38	0.39	48	93	32
9496	7.3	-42.269	172.429	10	140	0.19	30	40	14
9509	7.1	-42.953	172.814	10	66	0.35	65	208	71
9538	5.3	-43.498	172.551	16	4	0.98	71	145	49
9573	5.2	-43.491	172.573	4	3	1.11	92	139	47
9585	6.5	-42.990	172.045	17	73	0.75	57	210	71
9648	7.2	-43.233	171.771	10	73	0.17	26	48	16
9814	6.1	-43.745	172.214	11	39	0.68	52	149	51
9919	6.9	-42.701	172.084	14	100	0.20	28	40	14
9943	6.4	-43.661	172.294	1	28	0.75	82	203	69
10030	5.1	-43.426	172.543	18	11	0.31	29	5	2
10335	5.7	-43.633	172.497	10	14	1.48	136	205	70

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
10506	7.5	-43.317	172.547	11	23	0.37	84	218	74
10515	6.0	-43.731	172.469	7	25	0.60	68	123	42
10586	7.5	-43.317	172.547	11	23	1.14	151	223	76
10735	7.1	-42.953	172.814	10	66	0.44	48	138	47
10871	6.3	-43.487	172.424	5	13	1.33	107	214	73
10946	5.2	-43.476	172.513	8	8	0.43	55	28	10
10969	6.9	-42.994	172.170	18	68	0.79	72	216	73
10972	6.3	-43.539	172.053	19	43	0.44	37	87	30
11029	6.0	-43.509	172.848	17	21	0.61	72	51	17
11171	6.8	-43.315	172.467	10	25	0.48	62	196	67
11184	6.1	-43.322	172.069	17	47	0.37	41	122	41
11321	5.0	-43.404	172.662	17	14	0.46	24	21	7
11479	7.1	-42.881	172.503	10	71	0.25	29	62	21
11520	6.2	-43.788	173.262	14	62	0.41	63	96	33
11564	5.7	-43.623	172.589	8	11	0.37	51	53	18
11578	7.5	-43.317	172.547	11	23	0.69	102	218	74
11628	5.5	-43.691	172.691	4	21	1.19	62	97	33
11672	6.0	-43.825	172.589	30	34	0.88	49	29	10
11676	6.4	-43.274	173.220	33	58	0.19	38	12	4
11693	5.9	-43.520	172.368	9	17	1.12	106	224	76
11799	7.4	-42.602	172.459	11	103	0.20	48	78	27
11801	6.1	-43.776	172.125	12	46	0.75	47	161	55
11848	7.0	-43.267	171.469	16	94	0.56	57	186	63
11952	5.3	-43.589	172.762	12	16	0.67	46	63	21
12037	7.0	-43.106	172.237	40	54	0.29	92	155	53
12079	7.6	-42.936	173.691	11	111	0.38	67	152	52
12112	7.0	-43.601	173.014	10	36	0.67	119	225	77
12140	7.1	-42.693	172.253	10	96	0.15	25	8	3

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
12178	5.7	-43.544	172.316	7	22	1.00	67	94	32
12230	7.0	-43.548	172.430	11	13	0.91	231	236	80
12259	5.6	-43.511	172.553	17	3	1.13	90	185	63
12369	6.4	-43.327	172.572	12	22	0.70	73	229	78
12418	7.2	-43.233	171.771	10	73	0.38	56	122	41
12495	5.1	-43.533	172.712	19	10	0.38	30	13	4
12796	6.9	-43.735	172.634	7	24	0.81	181	176	60
12913	7.2	-42.823	172.917	10	82	0.33	44	182	62
12935	5.8	-43.383	172.424	13	20	0.47	67	152	52
12937	7.5	-42.302	173.664	10	162	0.18	37	89	30
13092	5.0	-43.571	172.539	5	7	0.50	59	72	24
13127	5.0	-43.516	172.521	2	5	0.68	110	59	20
13231	5.1	-43.547	172.829	14	20	0.93	59	75	26
13354	6.7	-43.059	172.567	33	51	0.53	60	174	59
13379	7.6	-42.936	173.691	11	111	0.29	72	217	74
13428	6.3	-43.708	172.490	2	22	0.38	50	117	40
13508	6.3	-43.415	172.567	23	12	0.54	63	139	47
13522	5.6	-43.556	172.509	12	7	0.43	68	62	21
13525	6.4	-43.319	172.569	17	22	0.54	86	222	76
13603	6.7	-43.608	172.506	20	11	0.77	173	234	80
13624	5.2	-43.478	172.728	4	13	1.39	95	83	28
13734	6.7	-43.655	172.685	5	17	3.68	297	292	99
13737	6.4	-43.090	172.275	30	54	0.30	32	35	12
13770	7.4	-43.317	171.663	11	78	0.27	39	141	48
13815	7.6	-42.936	173.691	11	111	0.24	32	51	17
13915	7.3	-42.793	171.632	11	112	0.16	31	48	16
13984	5.7	-43.387	172.513	8	16	0.61	54	90	31
14121	7.1	-43.490	171.992	10	48	0.50	72	213	72

Event ID	Magnitude	Latitude	Longitude	Depth	Distance[2]	Max PGA	Max PGV	Liquefied sites (1km grid)	
	M_W	°	°	km	km	g	cm/s	No	%
14158	5.6	-43.524	172.409	16	14	0.59	77	130	44
14163	5.7	-43.452	172.466	1	12	0.76	153	93	32
14250	5.2	-43.470	172.500	13	9	0.56	67	85	29
14262	6.6	-43.521	172.359	19	18	0.35	50	159	54
14271	7.2	-43.380	171.838	10	62	0.35	45	123	42
14348	5.6	-43.378	173.000	3	37	0.44	53	64	22
14403	5.0	-43.620	172.725	29	16	0.28	25	3	1

System performance

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
36	3	1	0	0	7.4	144	2	8	5	13	8	39.6	39.6
74	0	0	1	3	8.6	112	2	3	2	8	5	44.6	45.5
164	1	0	1	3	4.0	151	2	3	2	3	2	39.7	39.7
170	14	6	5	16	31.0	203	3	22	14	49	32	57.2	67.5
236	2	1	1	3	9.5	130	2	11	7	12	8	39.9	42.4
320	39	16	10	31	68.0	472	8	67	44	94	61	79.5	88.1
355	3	1	2	6	17.5	163	3	5	3	22	14	44.3	48.0
459	11	5	6	19	32.9	186	3	19	12	46	30	40.8	47.8
479	0	0	1	3	8.6	72	1	3	2	8	5	17.7	18.9
633	11	5	3	9	25.6	153	2	18	12	38	25	34.7	39.6
869	7	3	3	9	21.7	103	2	10	6	21	14	43.3	47.1
1085	0	0	1	3	8.1	137	2	16	10	25	16	42.3	42.9
1134	7	3	6	19	23.8	158	3	16	10	24	16	39.1	41.8
1202	12	5	3	9	12.6	257	4	23	15	31	20	70.5	72.7
1213	0	0	1	3	3.4	63	1	1	1	1	1	22.9	22.9
1263	33	14	7	22	63.8	471	8	55	36	73	47	79.2	86.4
1304	1	0	1	3	8.1	82	1	3	2	15	10	20.9	22.1
1426	0	0	1	3	1.7	52	1	1	1	2	1	22.9	23.8
1428	0	0	1	3	0.4	66	1	1	1	1	1	24.3	24.3
1468	38	16	8	25	58.6	355	6	37	24	70	45	71.6	79.4
1516	1	0	1	3	1.7	104	2	5	3	6	4	30.1	30.1
1537	27	11	5	16	26.7	295	5	23	15	38	25	65.9	72.0
1603	20	8	3	9	39.7	243	4	43	28	57	37	50.3	58.0
1611	1	0	4	13	17.8	92	1	11	7	27	18	29.6	33.7
1628	3	1	1	3	8.6	152	2	18	12	21	14	28.2	30.4

[3] Analysis excluding interdependency
[4] Analysis including interdependency

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
1636	3	1	2	6	10.6	95	2	11	7	18	12	28.1	29.4
1645	0	0	1	3	3.6	38	1	1	1	4	3	10.8	13.7
1672	2	1	4	13	35.6	180	3	17	11	26	17	51.0	52.9
1721	22	9	6	19	25.0	280	5	29	19	44	29	51.8	57.8
1789	1	0	2	6	9.1	154	2	11	7	24	16	51.0	51.2
1797	0	0	1	3	8.1	114	2	9	6	17	11	36.7	37.6
1802	7	3	2	6	4.3	182	3	7	5	10	6	44.3	46.3
1835	0	0	1	3	8.1	97	2	2	1	15	10	33.4	34.8
1837	5	2	5	16	34.7	167	3	9	6	42	27	39.1	45.2
1935	2	1	1	3	2.0	166	3	10	6	10	6	42.8	42.8
1953	0	0	1	3	8.6	112	2	3	2	8	5	34.9	36.2
1987	4	2	2	6	21.3	171	3	23	15	31	20	38.8	39.4
2213	13	5	6	19	36.9	244	4	33	21	40	26	58.4	61.9
2241	1	0	1	3	9.2	56	1	2	1	5	3	30.5	33.7
2245	0	0	1	3	0.6	95	2	3	2	3	2	31.5	31.5
2248	0	0	1	3	0.6	65	1	6	4	6	4	29.2	29.2
2279	11	5	5	16	26.4	132	2	6	4	21	14	35.5	38.4
2281	3	1	1	3	10.7	96	2	6	4	24	16	38.5	44.2
2285	11	5	7	22	35.5	271	4	30	19	51	33	61.6	64.6
2291	0	0	1	3	8.1	51	1	2	1	14	9	22.3	24.4
2320	27	11	8	25	50.5	323	5	37	24	60	39	64.7	73.3
2428	3	1	7	22	37.8	159	3	13	8	29	19	54.0	62.5
2507	5	2	2	6	14.6	148	2	10	6	14	9	41.1	43.5
2651	0	0	1	3	9.6	49	1	4	3	11	7	24.5	24.8
2733	6	2	6	19	21.5	132	2	16	10	24	16	34.2	35.3
2782	40	16	11	34	50.3	506	8	45	29	72	47	76.6	82.8
2785	20	8	5	16	44.4	250	4	27	18	41	27	61.6	67.0
2807	0	0	1	3	9.6	67	1	3	2	9	6	18.6	23.2

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
2886	3	1	2	6	18.4	141	2	9	6	31	20	46.2	51.9
2916	4	2	3	9	19.8	248	4	22	14	31	20	63.4	66.6
3029	1	0	2	6	10.7	130	2	17	11	28	18	31.8	34.7
3169	0	0	1	3	8.1	39	1	3	2	16	10	14.9	16.5
3233	1	0	1	3	10.4	111	2	3	2	9	6	44.7	44.7
3527	59	24	14	44	81.3	579	9	84	55	112	73	90.1	94.9
3688	0	0	1	3	10.4	80	1	3	2	9	6	31.1	31.1
3746	1	0	2	6	2.5	61	1	3	2	3	2	30.1	30.1
3754	1	0	1	3	16.6	150	2	16	10	23	15	47.9	50.1
3805	12	5	4	13	20.0	213	3	14	9	25	16	51.2	56.6
3809	8	3	3	9	23.0	182	3	11	7	21	14	49.7	50.7
3848	14	6	5	16	27.9	192	3	16	10	28	18	42.7	47.2
3855	0	0	1	3	8.1	64	1	0	0	14	9	21.2	23.5
3871	19	8	3	9	40.9	223	4	30	19	50	32	49.2	60.2
3932	0	0	1	3	8.1	47	1	1	1	14	9	14.5	17.6
3941	8	3	2	6	19.3	145	2	12	8	28	18	35.8	40.2
3997	5	2	2	6	5.2	202	3	18	12	18	12	52.2	52.2
4086	0	0	2	6	17.4	79	1	4	3	25	16	23.1	28.5
4216	24	10	6	19	43.2	321	5	30	19	47	31	65.2	70.8
4225	20	8	7	22	45.3	290	5	29	19	71	46	56.4	65.0
4353	40	16	9	28	51.7	391	6	56	36	97	63	72.1	78.0
4360	2	1	1	3	10.4	73	1	4	3	9	6	34.2	34.2
4373	5	2	2	6	17.4	151	2	15	10	27	18	33.0	36.5
4428	0	0	4	13	28.3	134	2	9	6	33	21	43.2	49.1
4433	0	0	1	3	9.6	73	1	4	3	10	6	24.2	26.9
4492	17	7	7	22	29.3	221	4	13	8	36	23	49.0	53.8
4496	48	20	12	38	66.8	585	9	64	42	84	55	81.9	86.2
4563	0	0	1	3	7.4	84	1	3	2	19	12	36.8	42.2

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
4678	23	9	6	19	29.2	301	5	36	23	54	35	70.9	72.7
4741	0	0	1	3	10.7	54	1	2	1	19	12	26.2	26.9
4823	22	9	9	28	45.0	383	6	51	33	95	62	60.4	64.1
4904	0	0	1	3	0.4	46	1	3	2	3	2	20.1	20.1
4944	71	29	14	44	75.1	549	9	75	49	103	67	86.6	91.8
4948	35	14	8	25	47.5	396	6	56	36	110	71	71.3	73.4
4969	12	5	6	19	48.7	262	4	26	17	69	45	58.4	66.5
4997	11	5	1	3	10.7	149	2	4	3	12	8	34.4	37.5
5030	0	0	1	3	17.9	48	1	0	0	24	16	22.2	26.8
5033	14	6	6	19	49.2	303	5	33	21	57	37	74.7	77.5
5055	7	3	3	9	10.6	119	2	16	10	22	14	32.0	33.0
5136	0	0	1	3	13.0	34	1	0	0	8	5	17.8	23.0
5240	7	3	2	6	19.5	205	3	20	13	45	29	50.6	57.2
5293	0	0	1	3	8.6	35	1	1	1	6	4	17.9	19.1
5308	0	0	3	9	6.9	165	3	9	6	11	7	42.6	42.6
5413	6	2	1	3	11.1	102	2	9	6	14	9	26.0	26.9
5581	8	3	2	6	26.6	217	4	21	14	30	19	56.0	59.2
5687	3	1	2	6	8.6	73	1	2	1	15	10	29.2	30.4
5767	2	1	2	6	17.5	113	2	8	5	17	11	34.1	34.8
5807	5	2	1	3	22.5	128	2	8	5	18	12	43.0	48.4
5816	7	3	2	6	5.9	169	3	13	8	16	10	48.7	50.7
6107	4	2	4	13	19.5	157	3	10	6	39	25	41.8	50.1
6273	15	6	6	19	38.8	270	4	38	25	55	36	55.0	60.0
6305	6	2	2	6	12.8	172	3	17	11	25	16	45.0	50.0
6385	0	0	3	9	12.6	149	2	8	5	15	10	38.0	39.9
6470	1	0	3	9	10.9	90	1	8	5	12	8	29.5	31.1
6519	7	3	2	6	28.4	151	2	18	12	33	21	33.8	38.6
6538	2	1	1	3	2.0	108	2	8	5	13	8	33.7	36.2

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
6547	0	0	1	3	3.6	41	1	0	0	3	2	12.9	15.9
6635	2	1	1	3	10.4	133	2	6	4	10	6	42.5	42.5
6739	6	2	3	9	30.5	201	3	15	10	45	29	49.3	57.6
6806	0	0	1	3	1.7	52	1	2	1	3	2	23.5	23.6
6817	0	0	1	3	0.6	40	1	0	0	0	0	28.6	28.6
6818	2	1	1	3	8.1	136	2	13	8	25	16	24.6	26.6
6865	0	0	1	3	22.5	49	1	1	1	14	9	19.2	22.8
7024	64	26	15	47	82.2	548	9	93	60	132	86	88.1	94.1
7056	30	12	4	13	44.2	339	6	37	24	57	37	78.4	82.5
7107	1	0	1	3	8.1	101	2	10	6	21	14	25.5	28.1
7132	42	17	10	31	64.9	363	6	42	27	80	52	73.4	83.4
7165	12	5	2	6	17.3	129	2	5	3	24	16	41.0	46.8
7189	0	0	2	6	2.3	55	1	1	1	1	1	19.9	19.9
7227	3	1	1	3	8.6	89	1	15	10	16	10	28.8	29.7
7269	4	2	4	13	15.9	186	3	22	14	27	18	41.4	44.9
7287	7	3	6	19	45.9	150	2	11	7	43	28	47.7	55.1
7294	6	2	3	9	5.7	108	2	10	6	11	7	30.6	30.6
7331	0	0	1	3	8.6	56	1	4	3	8	5	23.8	24.7
7347	4	2	1	3	17.9	225	4	13	8	28	18	46.3	52.2
7506	40	16	15	47	75.3	473	8	47	31	77	50	75.5	81.9
7514	0	0	1	3	9.6	51	1	2	1	7	5	22.7	28.0
7567	0	0	2	6	8.6	109	2	11	7	23	15	35.7	38.2
7589	0	0	1	3	9.2	73	1	7	5	11	7	22.4	25.6
7688	0	0	1	3	0.4	78	1	2	1	2	1	22.9	22.9
7785	5	2	4	13	31.9	170	3	19	12	44	29	45.0	50.6
7873	21	9	11	34	65.9	326	5	40	26	78	51	62.8	74.2
7902	3	1	3	9	18.2	144	2	8	5	15	10	45.3	49.4
8045	3	1	4	13	39.2	206	3	17	11	42	27	57.0	63.7

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
8187	3	1	0	0	9.2	180	3	17	11	20	13	42.7	45.1
8384	1	0	1	3	4.0	148	2	8	5	8	5	45.8	45.8
8445	0	0	2	6	3.3	41	1	4	3	5	3	22.5	22.7
8458	0	0	1	3	3.6	63	1	5	3	7	5	23.6	25.6
8524	0	0	1	3	8.1	90	1	4	3	14	9	22.1	23.5
8552	0	0	1	3	0.4	61	1	5	3	5	3	18.7	18.7
8716	13	5	4	13	18.2	159	3	21	14	33	21	46.5	49.7
8839	1	0	1	3	10.7	122	2	6	4	24	16	40.0	43.0
8983	2	1	1	3	21.3	69	1	3	2	16	10	31.4	32.3
9158	0	0	1	3	8.6	59	1	4	3	7	5	20.9	21.8
9188	0	0	2	6	12.8	67	1	0	0	7	5	20.7	25.8
9196	37	15	13	41	78.4	386	6	55	36	95	62	74.7	86.8
9239	9	4	2	6	17.9	139	2	23	15	32	21	31.2	32.8
9326	0	0	2	6	12.9	75	1	5	3	10	6	25.1	25.1
9404	12	5	5	16	22.9	224	4	20	13	23	15	54.8	55.6
9408	69	28	23	72	100.0	771	13	98	64	142	92	93.5	98.6
9495	4	2	1	3	21.3	102	2	4	3	17	11	35.4	38.7
9496	0	0	1	3	0.4	65	1	2	1	2	1	18.9	18.9
9509	8	3	4	13	15.3	249	4	15	10	18	12	58.3	61.5
9538	7	3	2	6	15.3	161	3	19	12	23	15	52.9	54.3
9573	6	2	4	13	35.4	209	3	25	16	34	22	50.2	54.2
9585	13	5	9	28	48.2	268	4	30	19	52	34	65.0	71.8
9648	1	0	1	3	9.2	76	1	3	2	7	5	36.5	40.4
9814	6	2	3	9	19.3	222	4	23	15	34	22	51.6	55.4
9919	0	0	1	3	13.0	69	1	1	1	9	6	23.1	24.6
9943	12	5	7	22	46.5	253	4	36	23	72	47	65.8	75.2
10030	0	0	1	3	8.6	47	1	0	0	5	3	16.8	17.7
10335	34	14	11	34	64.9	304	5	41	27	71	46	71.4	77.1

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
10506	21	9	7	22	31.0	317	5	38	25	66	43	66.5	77.8
10515	8	3	3	9	20.2	204	3	18	12	34	22	53.3	57.1
10586	36	15	15	47	75.2	420	7	57	37	90	58	83.9	90.0
10735	6	2	3	9	12.1	222	4	22	14	35	23	47.2	54.0
10871	29	12	13	41	70.9	359	6	44	29	111	72	68.2	76.3
10946	1	0	1	3	0.4	73	1	4	3	4	3	32.4	32.4
10969	52	21	17	53	78.4	427	7	63	41	93	60	80.2	85.9
10972	17	7	2	6	12.6	153	2	3	2	11	7	40.2	43.5
11029	1	0	2	6	17.4	100	2	11	7	30	19	22.2	26.3
11171	5	2	2	6	11.7	210	3	16	10	26	17	56.9	61.6
11184	1	0	1	3	10.4	132	2	10	6	14	9	45.0	45.0
11321	0	0	1	3	8.6	55	1	3	2	7	5	21.9	23.8
11479	1	0	1	3	4.0	116	2	8	5	8	5	30.7	30.7
11520	7	3	1	3	2.0	131	2	14	9	14	9	40.6	40.6
11564	0	0	2	6	10.3	74	1	4	3	12	8	32.8	37.2
11578	22	9	8	25	66.7	391	6	50	32	107	69	74.1	85.8
11628	17	7	3	9	19.4	186	3	27	18	33	21	49.8	49.8
11672	3	1	2	6	10.5	72	1	6	4	18	12	22.5	25.8
11676	0	0	1	3	4.0	44	1	3	2	3	2	14.9	14.9
11693	62	25	17	53	100.0	459	7	74	48	135	88	86.5	90.6
11799	0	0	1	3	4.0	84	1	3	2	3	2	24.1	24.1
11801	0	0	1	3	10.1	167	3	13	8	23	15	40.7	45.2
11848	0	0	3	9	20.3	206	3	17	11	25	16	53.4	56.8
11952	11	5	5	16	27.6	133	2	13	8	26	17	38.5	40.0
12037	0	0	1	3	0.4	141	2	7	5	7	5	35.0	35.0
12079	12	5	2	6	22.7	251	4	19	12	29	19	56.4	56.6
12112	36	15	8	25	56.5	365	6	44	29	66	43	69.4	73.6
12140	0	0	1	3	9.2	42	1	1	1	5	3	13.5	14.2

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
12178	1	0	3	9	29.1	109	2	2	1	17	11	39.6	45.6
12230	46	19	16	50	87.6	646	10	88	57	132	86	95.2	96.2
12259	33	14	7	22	67.6	307	5	47	31	83	54	64.2	77.0
12369	21	9	5	16	34.3	345	6	49	32	62	40	68.1	74.2
12418	15	6	5	16	34.7	171	3	18	12	33	21	48.3	57.4
12495	0	0	1	3	1.7	38	1	0	0	1	1	15.0	15.0
12796	16	7	1	3	14.5	237	4	24	16	41	27	64.6	67.4
12913	6	2	1	3	20.7	189	3	10	6	23	15	48.7	57.1
12935	5	2	1	3	15.4	173	3	10	6	28	18	48.1	51.0
12937	0	0	2	6	18.3	96	2	3	2	15	10	35.4	38.3
13092	1	0	1	3	9.2	93	2	9	6	11	7	36.2	39.4
13127	3	1	0	0	9.2	101	2	5	3	9	6	28.9	32.0
13231	3	1	3	9	21.9	134	2	18	12	29	19	31.0	33.3
13354	6	2	0	0	9.2	180	3	8	5	12	8	46.2	49.4
13379	7	3	3	9	19.1	255	4	24	16	40	26	55.8	58.9
13428	2	1	3	9	18.4	147	2	17	11	19	12	44.4	45.8
13508	13	5	6	19	27.6	240	4	29	19	41	27	56.8	58.7
13522	2	1	2	6	0.7	99	2	12	8	12	8	24.8	24.8
13525	29	12	11	34	69.3	361	6	48	31	92	60	73.1	86.2
13603	31	13	7	22	58.8	407	7	51	33	87	56	73.8	81.2
13624	14	6	2	6	26.9	179	3	26	17	37	24	38.0	41.8
13734	72	30	24	75	100.0	1181	19	129	84	144	94	96.2	97.9
13737	0	0	1	3	0.6	80	1	0	0	0	0	19.4	19.4
13770	2	1	2	6	11.9	123	2	3	2	8	5	31.4	32.3
13815	0	0	1	3	0.4	56	1	2	1	2	1	24.1	24.1
13915	0	0	1	3	10.1	60	1	1	1	14	9	23.5	25.7
13984	2	1	2	6	10.3	124	2	7	5	15	10	34.9	38.4

Event ID	Cable failures		Substation failures		EPN performance	Pipe failures		Pumping station failures				WSS performance	
								exc.[3]		inc.[4]		exc.[3]	inc.[4]
	No	%	No	%	%	No	%	No	%	No	%	%	%
14121	26	11	6	19	53.9	331	5	33	21	70	45	75.8	81.6
14158	13	5	5	16	32.1	142	2	20	13	35	23	49.2	56.8
14163	0	0	1	3	8.1	129	2	13	8	23	15	30.8	32.2
14250	3	1	3	9	9.9	96	2	19	12	23	15	36.0	36.7
14262	3	1	1	3	0.4	166	3	5	3	5	3	41.4	41.4
14271	1	0	2	6	5.8	159	3	5	3	5	3	47.0	47.0
14348	9	4	2	6	0.7	117	2	13	8	13	8	27.0	27.0
14403	0	0	1	3	0.4	37	1	0	0	0	0	17.3	17.3